

Result No.	Query	Score	Query			ID	Description
			Match	Length	DB		
1	1998	100.0	456	18	AAW36050	Hybrid Marek's dis	
2	1994	99.8	1086	15	AAW36051	Hybrid Marek's dis	
3	1910	95.6	615	18	AAW63230	Mycoplasma gallise	
4	1856	92.9	610	15	AAW63229	Mycoplasma gallise	
5	1812	80.7	368	14	AAW44493	Mycoplasma gallise	
6	1612	80.7	368	15	AAW44327	Mycoplasma gallise	
7	1604	80.3	368	16	AAW76955	Mycoplasma gallise	
8	1142	57.2	235	10	AAW93646	Amino acid sequenc	
9	1142	57.2	235	11	AAW05081	MG-1 antigen. AA	
10	1142	57.2	261	11	AAW05082	TMG-1 antigen. A	
11	1142	57.2	261	16	AAW79911	M.gallisepticum 26	

DR WPI; 1997-503046/46.
 XX N-PSDB; AAT96595.
 XX Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 PS Disclosure; Page 16-19; 51pp; Japanese.
 XX
 XX This sequence represents the chimeric protein 40 K-S which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gp fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 XX Sequence 456 AA;
 SQ
 Query Match 100.0%; Score 1998; DB 18; Length 456;
 Best Local Similarity 100.0%; Pred. No. 2.1e-126;
 Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CMSITKDNPNNGQQLAARMEITDLINAKAMTILASLDYAKIEASISSAYSEAEVYN 60
 DB 64 cmsitkddanpnngqqlaarmeltldlinakamtllasiqdyakieasissayseaeavn 123
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 120
 DB 124 nlnnatleqlkmaaktulesainganctdktfdnehpnlveaykalkttleqatnleqls 183
 QY 121 STAYNIRNLVDLYNKASLIKTLDPLNGGTLDSNEITTANKNINNTLSTINEQKTN 180
 DB 184 staynqirnlvlynkassliktldpninggtldgsneittankninnltstineqktn 243
 QY 181 ADALSNFIFKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 240
 DB 244 adalsnfikkvionneqsfvgftftnanvqpsnysvafsadvtvnykyartvwnge 303
 QY 241 PSSRILANTNSITDVSWIYSLAGTNTKYQFSFNGYGPSTGYLYFPYKLVKAADANNVGLQ 300
 DB 304 pssrilantnsitdsvwiyslagntkyqfsfngygpstgylyfpykvlkaadannvqlg 363
 QY 301 YKLNGNVOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 360
 DB 364 yklngnvqqvefatstsannttanptpavdeikvakivlsglrfgqntielsvptgegn 423
 QY 361 MNKVAPMIGNIYLSNENNAADKIPGYRRPCTFL 393
 DB 424 mnkvapmigniylyssnennadkpggyrrpctfl 456
 RESULT 2
 AAW36051
 ID AAW36051 standard; Protein; 1086 AA.
 XX
 XX AAW36051;
 XX
 XX 15-JUL-1998 (first entry)
 XX
 DE Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.
 XX
 XX Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
 KW antigen; vaccine; poultry.
 XX
 OS Chimeric - Marek's disease gammaherpesvirus.
 OS Chimeric - Mycoplasma gallisepticum.
 XX
 XX Key Location/Qualifiers
 FT 1..672
 FT Region /note= "derived from Marek's disease virus gp protein"
 FT 593..1086
 FT Region /note= "derived from M. gallisepticum antigen"
 FT

XX WO9736924-A1.
 XX
 PD 09-OCT-1997.
 XX
 PF 28-MAR-1997; 97WO-JP01084.
 XX
 PR 29-MAR-1996; 96JP-0103548.
 XX
 XX (JAPG) NIPPON ZEON KK.
 XX
 PI Saito S, Tsuzaki Y, Yanagida N;
 XX
 XX WPI; 1997-503046/46.
 DR
 DR N-PSDB; AAT96596.
 XX
 XX Fusion protein comprising herpes virus outer membrane protein and
 PT antigenic polypeptide - for prevention of infection by Mycoplasma
 PT gallisepticum, especially in poultry
 XX
 PS Disclosure; Page 22-30; 51pp; Japanese.
 XX
 XX This sequence represents the chimeric protein 40 K-C which comprises a
 CC fragment of the Marek's disease virus outer membrane protein gp fused
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
 CC protein can be used in recombinant live vaccines for prevention of
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane
 CC protein shows antigenicity in poultry.
 XX
 XX Sequence 1086 AA;
 SQ
 Query Match 99.8%; Score 1994; DB 18; Length 1086;
 Best Local Similarity 99.7%; Pred. No. 1.2e-125;
 Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CMSITKDNPNNGQQLAARMEITDLINAKAMTILASLDYAKIEASISSAYSEAEVYN 60
 DB 694 cmsitkddanpnngqqlaarmeltldlinakamtllasiqdyakieasissayseaeavn 753
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 120
 DB 754 nlnnatleqlkmaaktulesainganctdktfdnehpnlveaykalkttleqatnleqls 813
 QY 121 STAYNIRNLVDLYNKASLIKTLDPLNGGTLDSNEITTANKNINNTLSTINEQKTN 180
 DB 814 staynqirnlvlynkassliktldpninggtldgsneittankninnltstineqktn 873
 QY 181 ADALSNFIFKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTVWNGDE 240
 DB 874 adalsnfikkvionneqsfvgftftnanvqpsnysvafsadvtvnykyartvwnge 933
 QY 241 PSSRILANTNSITDVSWIYSLAGTNTKYQFSFNGYGPSTGYLYFPYKLVKAADANNVGLQ 300
 DB 934 pssrilantnsitdsvwiyslagntkyqfsfngygpstgylyfpykvlkaadannvqlg 993
 QY 301 YKLNGNVOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 360
 DB 994 yklngnvqqvefatstsannttanptpavdeikvakivlsglrfgqntielsvptgegn 1053
 QY 361 MNKVAPMIGNIYLSNENNAADKIPGYRRPCTFL 393
 DB 1054 mnkvapmigniylyssnennadkpggyrrpctfl 1086
 RESULT 3
 AAR63230
 ID AAR63230 standard; Protein; 615 AA.
 XX
 XX AAR63230;
 XX
 XX 23-JUN-1995 (first entry)
 XX

```

DE Mycoplasma gallisepticum antigen (UM-67).
KW recombinant avipox virus; live vaccine; mycoplasma antigen.
XX Mycoplasma gallisepticum.
OS
XX
XX Key Location/Qualifiers
XX Protein 1..615
XX /note= "Trp residues correspond to TGA codons"
XX
XX WO9423019-A.
XX 13-OCT-1994.
XX
XX 31-MAR-1994; 94WO-JP00541.
XX
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX
XX (JAPG ) NIPPON ZEON KK.
XX (SHIO ) SHIONOGI & CO LTD.
XX
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX Saeki S, Saitos, Takahashi K;
XX
XX WPI; 1994-333181/41.
XX N-PSDB; AAQ77857.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX
XX Claim 4; Page 87-91; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX pUM-67 containing an open reading frame was sequenced (AAQ77857). The
XX ORF encodes an antigenic polypeptide (AAR63230). A recombinant avipox
XX virus comprising the coding sequence can be used as a live vaccine to
XX protect against infection by Mycoplasma gallisepticum.
XX
XX Sequence 615 AA;

Query Match 95.6%; Score 1910; DB 15; Length 615;
Best Local Similarity 98.2%; Pred. No. 2.5e-120;
Matches 376; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CMSITKDDANPNNGQTGLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 60
Db 27 cmsitkddanpnngqtlqaarmeltdlinakartlasldyakielasaysaeatvn 86

Qy 61 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
Db 87 nlnlatleqlkmaktnlesainqantdkttfdnehpnlveaykalkttleqratnleqla 146

Qy 121 STAYNOIRNNLVLDYNKASSLITTKTLDPLNGTLLDSNEITTANKNINNTLSTINEQKTN 180
Db 147 staynqirnnlvldynkasslittktdplnggmlldsneltvnrlnintstineqktn 206

Qy 181 ADALSNSFIKKVIONNEQSFYGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTWNGDE 240
Db 207 adalnsfikkvionneqsfygtftnanvqpsnysfvafsadvtvpnykyarrtwngde 266

Qy 241 PSSRLANTNSITDVSWYISLAGNTKYQFSFNGYSTGYLYPPYKLVKAADANNVGLQ 300
Db 267 pssrlantnsitdvswwyislagntkyqfsfngystgylyfpykvlvkaadannvqlq 326

Qy 301 YKLNNGNVQVEEATSTSANNTTANPTPAVDKIKVAKIVLSGLRFGONTIELSVPTGEGN 360
Db 327 yklnngnvqqvefatustsannttanptpavdeikvakivlsglrfgontielsvptgegn 386

Qy 361 MNKVAPMIGNIYLSNENNADKI 383
Db 383 mknvapmigniylyssennadki

```

```

Db 387 mnkvapmigniylyssennadki 409

RESULT 4
AAR63229 standard; Protein; 610 AA.
XX
XX AAR63229;
XX
XX 23-JUN-1995 (first entry)
XX
XX Mycoplasma gallisepticum antigen (UM-66).
XX recombinant avipox virus; live vaccine; mycoplasma antigen.
XX Mycoplasma gallisepticum.
XX
XX Key Location/Qualifiers
XX Protein 1..610
XX /note= "Trp residues correspond to TGA codons"
XX
XX WO9423019-A.
XX 13-OCT-1994.
XX
XX 31-MAR-1994; 94WO-JP00541.
XX
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX
XX (JAPG ) NIPPON ZEON KK.
XX (SHIO ) SHIONOGI & CO LTD.
XX
XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
XX Saeki S, Saitos, Takahashi K;
XX
XX WPI; 1994-333181/41.
XX N-PSDB; AAQ77856.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX
XX Claim 4; Page 78-81; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX pUM-66 containing an open reading frame was sequenced (AAQ77856). The
XX ORF encodes an antigenic polypeptide (AAR63229). A recombinant avipox
XX virus comprising the coding sequence can be used as a live vaccine to
XX protect against infection by Mycoplasma gallisepticum.
XX
XX Sequence 610 AA;

Query Match 92.9%; Score 1856; DB 15; Length 610;
Best Local Similarity 95.5%; Pred. No. 1.1e-116;
Matches 365; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CMSITKDDANPNNGQTGLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 60
Db 27 cmsitkddanpnngqtlqaarmeltdlinakartlasldyakielasaysaeatvn 86

Qy 61 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
Db 87 nlnlatleqlkmaktnlesainqantdkttfdnehpnlveaykalkttleqratnleqla 146

Qy 121 STAYNOIRNNLVLDYNKASSLITTKTLDPLNGTLLDSNEITTANKNINNTLSTINEQKTN 180
Db 147 staynqirnnlvldynkasslittktdplnggmlldsneltvnrlnintstineqktn 206

Qy 181 ADALSNSFIKKVIONNEQSFYGTFTNANVQPSNYSFVAFSADVTVPVNYKYARRTWNGDE 240
Db 207 adalnsfikkvionneqsfygtftnanvqpsnysfvafsadvtvpnykyarrtwngde 266

```

QY 241 PSSRILANTNSITDWSIYSLAGTNTKYQFSNYPSTGYLYFPYKLVKAADANNVGLQ 300
 DB 267 psrillantnsitdwsyislagtnktyqfsnypstgylyfpyklvkaadannvqlq 326
 QY 301 YKLNGNVOQVEFATSTANNTTANPTPAVDIKVAKIVLSGLRFGQNTIELSVPTGEGN 360
 DB 327 yklngnvqvefatstsannttanptpavdeikvakivlslgrfgqntielsvptgern 386
 QY 361 MNKVAPMIGNIYLSNENNAADK 382
 DB 387 mnkvapmignmytssnaeank 408
 RESULT 5
 AAR44493
 ID AAR44493 standard; Protein; 368 AA.
 XX
 AC AAR44493;
 XX
 DT 16-JUN-1994 (first entry)
 XX
 DE Mycoplasma gallisepticum 40kd antigen.
 XX
 KW Vaccine; mycoplasma infection; poultry; fowl.
 XX
 OS Mycoplasma gallisepticum.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 262 /note= "corresponds to NNN codon in AAQ53419"
 FT Misc-difference 283 /note= "corresponds to NNN codon in AAQ53419"
 FT
 FT
 XX WO9324646-A.
 PN
 XX
 XX
 PD 09-DEC-1993.
 PF
 PF 28-MAY-1993; 93WO-JP00715.
 XX
 XX 29-MAY-1992; 92JP-0138819.
 XX
 XX (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX
 XX Aoyama S, Fujisawa A, Iritani Y, Ohkawa S, Saito S;
 PI
 XX WPI; 1993-405837/50.
 DR N-PSDB; AAQ53419.
 XX
 XX Mycoplasma gallisepticum antigen and DNA coding for it - useful
 PT for vaccination of fowl against mycoplasma infections
 XX
 XX Claim 2; Page 23-26; 37pp; Japanese.
 PS
 XX The sequence coding for the 40kba antigen was obtained by PCR
 CC amplification of M.gallisepticum genomic DNA. The antigen reacts
 CC with Mycoplasma-immune or Mycoplasma-infected serum and can be used
 CC as a vaccine to protect fowl from M.gallisepticum infection.
 XX
 XX Sequence 368 AA;
 SQ
 Query Match 80.7%; Score 1612; DB 14; Length 368;
 Best Local Similarity 95.0%; Pred. No. 1.4e-100;
 Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
 QY 1 CMSITKDKANPNNGQTQLEARMELTDLINAKAMTSLASLDQYAKIEASLSAYSEAEVTN 60
 DB 27 cmsitkdkanpnngqtqlqaarmeltdlinakartlasldqyakieaslsayseaeatvn 86
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120

DB 87 nlnnatleqlkmaaktnlesainqantdkttfdnebpnlveaykalkttleqratnleqla 146
 QY 121 STAYNQIRNNLVDLNFKASSLITKTLDPNLGCTLLDSNEITANKNINNTLSTINEQKTN 180
 DB 147 staynqirnnlvdllynnassllitktldplnggmlldsnaitvnrnintstineqkcn 206
 QY 181 ADALNSFTKFKVIONNEQSFVGTETNANVOPSNYSFVAFSADVTVPVNYKYARRTVWNGDE 240
 DB 207 adalnsftkfkviqnnedsvgtetnannvqpsnysfvaifsadvtvpnykyarrtvxngde 266
 QY 241 PSSRILANTNSITDWSIYSLAGTNTKYQFSNYPSTGYLYFPYKLVKAADANNVGLQ 300
 DB 267 psrillantnsitdwsyislagtnktyqfsnypstgylyfpyklvkaadannvqlq 326
 QY 301 YKLNGNVOQVEFATSTANNTTANPTPAVDIKVAK 337
 DB 327 yklngnvqvefatstsannttanptqqlmrklk 363
 RESULT 6
 AAR63227
 ID AAR63227 standard; Protein; 368 AA.
 XX
 AC AAR63227;
 XX
 DT 23-JUN-1995 (first entry)
 XX
 DE Mycoplasma gallisepticum 40kd antigen.
 XX
 KW recombinant avipox virus; live vaccine; mycoplasma 40kd antigen;
 KW TTM-1.
 XX
 OS Mycoplasma gallisepticum.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 262 /note= "corresponds to a NNN codon"
 FT Misc-difference 283 /note= "corresponds to a NNN codon"
 FT
 FT
 XX WO9423019-A.
 PN
 XX
 XX 13-OCT-1994.
 XX
 XX 31-MAR-1994; 94WO-JP00541.
 XX
 XX 31-MAR-1993; 93JP-0074139.
 PR 30-SEP-1993; 93JP-0245625.
 XX
 XX (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX
 XX Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
 PI Saeki S, Saitos, Takahashi K;
 XX
 XX WPI; 1994-333181/41.
 DR N-PSDB; AAQ77854.
 XX
 XX Recombinant avipox virus combining DNA encoding a polypeptide -
 PT exhibiting antigenicity of mycoplasma, useful for the production
 PT of a live vaccine
 XX
 PS Claim 4; Page 71-74; 123pp; Japanese.
 XX
 XX The plasmid pUTM-1P contains a sequence (the TTM-1 gene) coding
 CC for the 40kd antigen of Mycoplasma gallisepticum under the control
 CC of a synthetic promoter. A 1300 bp restriction fragment containing
 CC the promoter-ORF sequence was excised and was used in the
 CC construction of a recombinant avipox virus vector comprising the
 CC TTM-1 gene, DNA encoding the signal membrane anchor peptide from
 CC Newcastle Disease Virus haemagglutinin neuraminidase and FVP
 CC sequences. The recombinant avipox virus is useful as a live vaccine

CC to protect against infection by Mycoplasma gallisepticum.

XX
SQ Sequence 368 AA;

Query Match 80.7%; Score 1612; DB 15; Length 368;
Best Local Similarity 95.0%; Pred. No. 1.4e-100;
Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 CMSITKDKANPNNGOTOLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 60
|||||
DB 27 cmsitkdkanpnngotqlqaarmeltdlinakartlasldyakiessaysaeactn 86
|||||
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 120
|||||
DB 87 nlnnatleqlkmaktlnlesainqantdkttfdnehpnlveaykalkttleqatnleqla 146
|||||
QY 121 STAYNQIRNNLVLYNKASSLITKTLDPNGTLLDSNEITTANKNNINTLSTINEOKTN 180
|||||
DB 147 staynqirnnlvlynnasslittkldplngmildsneittvnrnintlstineqktn 206
|||||
QY 181 ADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 240
|||||
DB 207 adalsnsfikkvionneqsfvgtftcnanvqpsnysfvafsadvtvnykyarrtvxngde 266
|||||
QY 241 PSSRLANTNSITDVSWIYSLAGTNTKYQFSNYGPGSTGYLYFPYKLVKAADANNVGLQ 300
|||||
DB 267 pssrlantnsitdvsxiyslagntkyqfsnygpgstgylyfpyklvkaadannvqlq 326
|||||
QY 301 YKLNGNVQVQVEFATSTANNTTANPTPAVDKIKVAK 337
|||||
DB 327 yklngnvqvqvefatstannnttanptqqlmrklk 363
|||||

RESULT 7

AAR76955
ID AAR76955 standard; Protein; 368 AA.

XX
AC AAR76955;

XX
DT 12-MAR-1996 (first entry)

XX
XX Mycoplasma gallisepticum antigenic protein TTM-1.

DE
XX Antigenic protein; vaccine; poultry; diagnosis; TTM-1.

XX
OS Mycoplasma gallisepticum.

XX
FH Key Location/Qualifiers

FT Misc-difference 262

FT /note= "any amino acid"

FT Misc-difference 283

FT /note= "any amino acid"

XX
PN JP07133295-A.

XX
PD 23-MAY-1995.

XX
PF 27-AUG-1993; 93JP-0213102.

XX
PR 27-AUG-1993; 93JP-0213102.

XX
PA (SHIO) SHIONOGI & CO LTD.

XX
XX WPI; 1995-220782/29.

DR
DR N-PSDB; AAQ94711.

XX
XX A new antigenic protein which reacts with Mycoplasma gallisepticum -

PT is useful in a component vaccine for use against poultry infected

PT with M. gallisepticum.

XX
XX Claim 6; Figs 5-6; 33pp; Japanese.

PS
XX

CC AAQ94711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein
TTM-1. TTM-1 can be used as a vaccine for M. gallisepticum
CC infectious diseases in poultry, and as a diagnostic agent for
CC M. gallisepticum.

XX
SQ Sequence 368 AA;

Query Match 80.3%; Score 1604; DB 16; Length 368;
Best Local Similarity 94.4%; Pred. No. 4.7e-100;
Matches 318; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 1 CMSITKDKANPNNGOTOLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 60
|||||
DB 27 cmsitkdkanpnngotqlqaarmeltdlinakartlasldyakiessaysaeactn 86
|||||
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 120
|||||
DB 87 nlnnatleqlkmaktlnlesainqantdkttfdnehpnlvqpykalkttleqatnleqla 146
|||||
QY 121 STAYNQIRNNLVLYNKASSLITKTLDPNGTLLDSNEITTANKNNINTLSTINEOKTN 180
|||||
DB 147 staynqirnnlvlynnasslittkldplngmildsneittvnrnintlstineqktn 206
|||||
QY 181 ADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 240
|||||
DB 207 adalsnsfikkvionneqsfvgtftcnanvqpsnysfvafsadvtvnykyarrtvxngde 266
|||||
QY 241 PSSRLANTNSITDVSWIYSLAGTNTKYQFSNYGPGSTGYLYFPYKLVKAADANNVGLQ 300
|||||
DB 267 pssrlantnsitdvsxiyslagntkyqfsnygpgstgylyfpyklvkaadannvqlq 326
|||||
QY 301 YKLNGNVQVQVEFATSTANNTTANPTPAVDKIKVAK 337
|||||
DB 327 yklngnvqvqvefatstannnttanptqqlmrklk 363
|||||

RESULT 8

AAP93646
ID AAP93646 standard; protein; 235 AA.

XX
AC AAP93646;

XX
DT 11-MAY-1990 (first entry)

XX
DE Amino acid sequence of Mycoplasma gallisepticum (MG1) polypeptide.

XX
KW Mycoplasma gallisepticum; Poultry vaccine; ss;

XX
OS Mycoplasma gallisepticum.

XX
PN EP345021-A.

XX
PD 06-DEC-1989.

XX
PF 31-MAY-1989; 89EP-0005441.

XX
PR 02-JUN-1988; 88JP-0136343.

XX
XX (JAPG) NIPPON ZEON KK.

PA (SHIO) SHIONOGI SEIYAKU KABUSHIKI KAISHA.

XX
PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;

XX
DR WPI; 1989-358393/49.

DR
DR N-PSDB; AAN92568.

XX
XX New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PT vaccines.

XX
PS Disclosure; Fig.1a; 31pp; English.

XX
XX This amino acid sequence of MG1 is encoded by M1 DNA and elicits an

CC antigen-antibody reaction with anti-MG poultry sera. It can be used as a
 CC vaccine to prevent and diagnose MG infection. Doseage is at least 1 micro
 CC gram vaccine / kg. No acute toxicity was noted with a dose of 5 mg / kg.
 XX
 SQ Sequence 235 AA;

Query Match 57.2%; Score 1142; DB 10; Length 235;
 Best Local Similarity 97.0%; Pred. No. 2.9e-69;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEAEVTN 60
 DB 1 cmsitkkdanpnngqtlgaarmeltdlinakartilasldqyakieaslsayseaeetvn 60
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
 DB 61 nlnnatleqlkmaaktulesainqantdkttfdnehpnlvaykalkttleqratnleгла 120
 QY 121 STAYNQIRNNLVLYNKASLIKTLDPLNGGTLSDNEITTANKNNINNTLSTINEQKTN 180
 DB 121 staynqirnnlvlynassliktldplnggmldneittvnrninntlstineqktn 180
 QY 181 ADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 235
 DB 181 adalsnsfikkvionneqsfvgtftnanvqpsnysfafsadvtpvnykyarritv 235

RESULT 9

AAR05081
 ID AAR05081 standard; protein; 235 AA.
 XX
 AC AAR05081;
 XX
 DT 08-OCT-1990 (first entry)
 XX
 DE MG-1 antigen.
 XX
 KW Mycoplasma gallisepticum; poultry; vaccine.
 XX

JP02111795-A.
 XX

24-APR-1990.
 XX

02-JUN-1989; 89JP-0136343.
 XX

02-JUN-1989; 89JP-0136343.
 XX

(JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI KK.
 XX

WPI; 1990-169109/22.
 DR N-PSDB; AAQ04686.
 XX

Diagnostic and vaccine for poultry mycoplasma serum - utilises
 antigen protein of the disease and recombinant vector
 incorporated with its coding gene.
 PT
 PS Claim 2; Fig 1a; 20pp; Japanese.
 XX

DNA encoding the protein can be inserted into an expression vector
 for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
 reaction with anti-mycoplasma gallisepticum poultry sera. It may also
 be ligated to other DNA to produce fusion proteins with an N-terminal
 bacterial enzyme sequence.
 CC
 CC See also AAR05081-2 and AAR06437-41.
 XX
 SQ Sequence 235 AA;

Query Match 57.2%; Score 1142; DB 11; Length 235;
 Best Local Similarity 97.0%; Pred. No. 2.9e-69;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEAEVTN 60
 DB 1 cmsitkkdanpnngqtlgaarmeltdlinakartilasldqyakieaslsayseaeetvn 60
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
 DB 61 nlnnatleqlkmaaktulesainqantdkttfdnehpnlvaykalkttleqratnleгла 120
 QY 121 STAYNQIRNNLVLYNKASLIKTLDPLNGGTLSDNEITTANKNNINNTLSTINEQKTN 180
 DB 121 staynqirnnlvlynassliktldplnggmldneittvnrninntlstineqktn 180
 QY 181 ADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 235
 DB 181 adalsnsfikkvionneqsfvgtftnanvqpsnysfafsadvtpvnykyarritv 235

RESULT 10

AAR05082
 ID AAR05082 standard; protein; 261 AA.
 XX
 AC AAR05082;
 XX
 DT 08-OCT-1990 (first entry)
 XX
 DE TMG-1 antigen.
 XX
 KW Mycoplasma gallisepticum; poultry; vaccine.
 XX

JP02111795-A.
 XX

24-APR-1990.
 XX

02-JUN-1989; 89JP-0136343.
 XX

02-JUN-1989; 89JP-0136343.
 XX

(JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI KK.
 XX

WPI; 1990-169109/22.
 DR N-PSDB; AAQ04687.
 XX

Diagnostic and vaccine for poultry mycoplasma serum - utilises
 antigen protein of the disease and recombinant vector
 incorporated with its coding gene.
 PT
 PS Claim 2; Fig 2; 20pp; Japanese.
 XX

DNA encoding the protein can be inserted into an expression vector
 for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
 reaction with anti-mycoplasma gallisepticum poultry sera. It may also
 be ligated to other DNA to produce fusion proteins with an N-terminal
 bacterial enzyme sequence.
 CC
 CC See also AAR05081 and AAR06437-41.
 XX
 SQ Sequence 261 AA;

Query Match 57.2%; Score 1142; DB 11; Length 261;
 Best Local Similarity 97.0%; Pred. No. 3.4e-69;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CMSITKKDANPNNGQTQLEAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEAEVTN 60
 DB 27 cmsitkkdanpnngqtlgaarmeltdlinakartilasldqyakieaslsayseaeetvn 86
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
 DB 87 nlnnatleqlkmaaktulesainqantdkttfdnehpnlvaykalkttleqratnleгла 146
 QY 121 STAYNQIRNNLVLYNKASLIKTLDPLNGGTLSDNEITTANKNNINNTLSTINEQKTN 180

Db 147 stayqirnnlvdlynnasslittkldplngmllsneittvnrnintstineqktn 206
 QY 181 ADALSNSEIKKVIQNEQSFVGTFTNANQPSNYSFVAFSADVTVPVNYKYARRTV 235
 Db 207 adalsnsfikkvigneqsfvgtftnanvqpsnysfvafsadvtvpvnykyarrtv 261

RESULT 11
 AAR79911
 ID AAR79911 standard; Protein; 261 AA.
 AC AAR79911;
 DT 19-JUL-1996 (first entry)
 DE M.gallisepticum 261 amino acid protein.
 KW Detection; probe; primer; PCR; amplification; secretion; lung;
 KW avian chronic respiratory disease; respiratory tract; nasal cavity.
 OS Mycoplasma gallisepticum.
 XX JP07236498-A.
 XX 12-SEP-1995.
 XX 25-FEB-1994; 94JP-0052764.
 XX 25-FEB-1994; 94JP-0052764.
 XX (JAPG) NIPPON ZEON KK.
 XX (SHIO) SHIONOGI & CO LTD.
 XX WPI; 1995-347462/45.
 XX N-PSDB; AAT04076.

Detection of Mycoplasma gallisepticum - for the quick detection,
 i.e. within one day, of avian chronic respiratory diseases
 Claim 3; Page 10-11; 11pp; Japanese.
 This is the amino acid sequence of a 261 amino acid protein encoded
 by a fragment of the Mycoplasma gallisepticum genome. The encoding
 sequence and the sequence of AAT04075 (encoding a 661 amino acid
 protein) can be used to detect M.gallisepticum using probes based on
 nucleotides 1125-1648 and primers based on nucleotides 449-466, the
 complement of bases 893-919, 1908-1934 and the complement of bases
 2184-2210 of AAT04075 and a probe based on nucleotides 718-41 of the DNA
 encoding this protein. The method using these sequences is faster i.e. is
 able to detect M.gallisepticum, which causes avian chronic respiratory
 diseases, within one day, from avian secretions, washings from the lung,
 respiratory tract, nasal cavity, etc.
 SQ Sequence 261 AA;

Query Match 57.28; Score 1142; DB 16; Length 261;
 Best Local Similarity 97.08; Pred. No. 3.4e-69;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CWSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSAETVN 60
 Db 27 cmsitckdanpnngqtlqarmeltdlinakartlasldyakyakleaslsayseae 86
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
 Db 87 nlnnatleqlkmaktlesainqantdkttfdnehpnlveaykalkttleqratnle 146
 QY 121 STAYNQIRNNLVDLYNKASSLITTKTLDPLNGTLLDSNEITTANKNNINTLSTINEQKTN 180
 Db 147 staynqirnnlvdlynnasslittkldplngmllsneittvnrnintstineqktn 206

QY 181 ADALSNSEIKKVIQNEQSFVGTFTNANQPSNYSFVAFSADVTVPVNYKYARRTV 235
 Db 207 adalsnsfikkvigneqsfvgtftnanvqpsnysfvafsadvtvpvnykyarrtv 261
 RESULT 12
 AAP93959
 ID AAP93959 standard; protein; 261 AA.
 XX AAP93959;
 AC AAP93959;
 DT 11-MAY-1990 (first entry)
 DE Amino acid (AA) sequence of TMG-1 polypeptide.
 KW Mycoplasma gallisepticum; Poultry vaccine; ss;
 OS Mycoplasma gallisepticum.
 XX EP345021-A.
 XX 06-DEC-1989.
 XX 31-MAY-1989; 89EP-0005441.
 XX 02-JUN-1988; 88JP-0136343.
 XX (JAPG) NIPPON ZEON KK; (SHIO) SHIONOGI SEIYAKU KABUSHIKI KAISHA.
 XX Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;
 WPI; 1989-358393/49.
 XX N-PSDB; AAN92574.
 XX New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
 PT vaccines.
 XX Disclosure; Fig.2; 31pp; English.
 CC This AA sequence of TMG-1 is encoded by TM-1 base sequence. It has the
 CC same sequence as that of a polypeptide expressed in Mycoplasma
 CC gallisepticum in nature. When the corresponding DNA sequence is inserted
 CC into a recombinant vector used to transform a host the antigen protein
 CC produced can be used as a vaccine to prevent and diagnose MG infection.
 XX Sequence 261 AA;
 SQ
 Query Match 55.9%; Score 1117; DB 10; Length 261;
 Best Local Similarity 96.2%; Pred. No. 1.6e-67;
 Matches 227; Conservative 3; Mismatches 4; Indels 2; Gaps 2;
 QY 1 CWSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSAETVN 60
 Db 27 cmsitckdanpnngqtlqarmeltdlinakartlasldyakyakleaslsayseae 86
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
 Db 87 nlnnatleqlkmaktlesainqantdkttfdnehpnlveaykalkttleqratnle 146
 QY 121 STAYNQIRNNLVDLYNKASSLITTKTLDPLNGTLLDSNEITTANKNNINTLSTINEQKTN 179
 Db 147 staynqirnnlvdlynnasslittkldplngmllsneittvnrn-intstineqk 205
 QY 180 NADALSNSEIKKVIQNEQSFVGTFTNANQPSNYSFVAFSADVTVPVNYKYARRTV 235
 Db 206 nadalsnsfikkvigneqsfvgtftnanvqpsnysfvafsadvtvpvnykyarrtv 261
 RESULT 13
 AAR63226
 ID AAR63226 standard; Protein; 661 AA.
 XX

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:46:22 ; Search time 29.99 seconds
(without alignments)
1259.190 Million cell updates/sec

Title: US-09-147-052-2_COPY_64_456

Perfect score: 1998
Sequence: 1 CMSITKKDANPNNGQTLEA.....SSNENNADKIPGYRRPGTEL 393

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	831.5	41.6	702	2 S48754	major surface prot
2	754	37.7	702	2 S48753	major surface prot
3	744	37.2	650	2 S48751	major surface prot
4	743	37.2	649	2 S48752	major surface prot
5	713	35.7	647	2 A49218	hemagglutinin homo
6	575.5	28.8	386	2 S48755	major surface prot
7	189.5	9.5	320	2 S51560	major surface prot
8	175	8.8	6713	2 B89921	hypothetical prote
9	173	8.7	1302	1 JC6009	surface-located me
10	165.5	8.3	661	2 Ac2442	hypothetical prote
11	164.5	8.2	2481	2 D90011	FatB protein [impo
12	163.5	8.2	3890	2 C89921	hypothetical prote
13	159	8.0	933	2 S41539	fibrogen-binding
14	156	7.8	807	2 B71605	hypothetical prote
15	155	7.8	1072	2 A86827	hypothetical prote
16	155	7.8	4688	2 F82885	hypothetical prote
17	154.5	7.7	682	2 S44131	subtilisin-like pr
18	154.5	7.7	1365	2 T30822	impl protein - Myc
19	151.5	7.6	1237	2 T71850	probable outer mem
20	151	7.6	624	2 PC6003	surface membrane p
21	150.5	7.5	1051	2 T18351	impl protein - Myc
22	145	7.3	2399	2 H71879	toxin-like outer m
23	144	7.2	1107	2 AC0976	probable autotrans
24	143.5	7.2	751	2 T40462	ser-lys rich hypot
25	142	7.1	135	2 B49218	hemagglutinin homo
26	142	7.1	2660	2 E85822	probable invasiv z
27	141.5	7.1	1645	2 F96907	phage-related prot
28	140.5	7.0	3194	2 D71917	toxin-like outer m
29	139.5	7.0	1487	2 AG2560	hypothetical prote

30 139.5 7.0 2401 2 T28676 rhoptry protein -
31 138.5 6.9 1314 1 TNBYR6 transcription regu
32 137.5 6.9 820 2 T17519 cell surface antiq
33 137.5 6.9 926 2 AE1130 conserved hypothet
34 137.5 6.9 989 2 D89852 fibrinogen-binding
35 137.5 6.9 1524 2 S68553 surface layer prot
36 137.5 6.9 4152 2 T31102 filamentous hemagg
37 137 6.9 719 2 S55119 hypothetical prote
38 136.5 6.8 568 2 E97066 membrane associate
39 136.5 6.8 1073 2 S14032 kinesin-related pr
40 136.5 6.8 2271 2 F90073 hypothetical prote
41 136 6.8 5005 2 F82884 hypothetical prote
42 135.5 6.8 1085 2 T38378 kinesin-like prote
43 135.5 6.8 1238 2 A64596 kinesin-like prote
44 135 6.8 2269 2 T28677 rhoptry protein -
45 134.5 6.7 589 2 B97806 hypothetical prote

ALIGNMENTS

RESULT 1

S48754

major surface protein (clone pMGAL.4) precursor - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum

C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999

C:Accession: S48754

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F

FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface

A:Reference number: S48751; MUID:95010739

A:Accession: S48754

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-702 <MAP>

A:Cross-references: EMBL:L28424

C:Genetics:

A:Genetic code: SGC3

A:Start codon: GTG

Query Match 41.6%; Score 831.5; DB 2; Length 702;
Best Local Similarity 42.0%; Pred. No. 3.2e-38;
Matches 173; Conservative 84; Mismatches 116; Indels 39; Gaps 8;

QY 8 DANPNG-----QTQLEARMELTDLINAKAMTILASLDYAKIEASLSAYSEATVNN 61
DB 68 NTPNGGGTDAQAOLAAAKKELSDLLATQNSLSTYADYANIONTLTAAYTTAKSTD 127
QY 62 NLNATLEQLKMAKTLESAINQANTDKTTFDNEHPNLVYAKKTTLEQRATNLEGLSS 121
DB 128 NTSATLEQVKSATSTLOTATDAASSKTSFDEKPNELIKAYNALKETLKKWRNLSGLTD 187
QY 122 TAYNQIRNRLVDLYNKASSLTKTLDPLNGGLLSNEITTTANKNNNTLSTINEQKTA 181
DB 188 SNFAIKTNLTALYSGGDKVTIKTLDPLM-GTAINLSAVSQANTNISNAVKLETWKTA 246
QY 182 DALNSFTKTKVIONNEQSPVGTFTNANVPNSYVFAESADVTP-----VNYKYARRTV 235
DB 247 TVLATSFVKEVLVKKNLGTIDT-TNNEQPGNYSFVGYSDVTTGSDNARNPNSFQRKV 305
QY 236 WNGD-----EPSSRLANTNSITDYSWIYSLAGTNTKYQFSFNSYGPSTGYLYFPYKLVK 290
DB 306 WTSNTDILSQPOAEGENQOAPDYSWIYNLTGMAKYSITFNYIGPSTGYLYFPYKLVN 365
QY 291 ADANNVGLQYKLNNGNVQOQVEFATS-----TSANN-----TTANPPPAVD 331
DB 366 SSDSDKVALEYKLNESAVKTIIDFSPSQTSPVASDATRENNRSTAAPAQGSTINPAPTLD 425
QY 332 ELKVAKIVLSGLRFQONTIELVSP-TGEGNMKNVAPMIGNIYLSNENNAKD 382
DB 426 DIKIAKVTLSNLKFGSNTIEFSVPTTAKEGTSKVAAPMIGNIYLTSSDRDVK 477

QY 8 DANPNNGO-----TQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTN 60
DB 48 DTNPGDGGGMMNAASQELAAARMGLTTFDSKAKNLGLYVDYKKTQNTLTTRAYDAAKTVL 107
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVAYKALKTTLEQRAATNLEGLS 120
DB 108 DNSSSTTQNLNEAKTRLETAIRTAATSQTDFEQHAELVKVYKELKTTLSNETATLAPYA 167
QY 121 STAYNQIRNNLDVLYNKASSLTITKLDPLNGTGLDLSNEITTANKNINNTL--STINQOK 178
DB 168 DAQYAGIKMHLISGLYDAGKAITTKLEPVEGDP--LTASAVMANTKIIEAIDKDEVLPQK 226
QY 179 TNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPV-----NKKY 230
DB 227 ENATKLADSFVKQVLVKEKITGVEEAAH--KAQPNYSFVGISVDITGTTGTGTSIPNWDY 285
QY 231 ARRTVW--NGDEPSSRILANT-----NSITDVSWIYSLAGTNTKYQFSFNSYGPSTGY 281
DB 286 AORTIFTNSDEP--RSISNTPADGOTMAQPLSNVSWIYSLAGTGAKYKTLFTFYGPSTGY 343
QY 282 LYFPYKLVKAADANNVGLQYKLNNGNVQOEFATSTSA-----NNTTANPTPAVDEIKVA 336
DB 344 LYFPYKLVNTSDQVKGLEYKLD-----ATKPSAITFGSDQTMNGKTPTVNDINVA 395
QY 337 KIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIGNYSNNENNADKI 383
DB 396 KVTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 437
RESULT 5
A:hemagglutinin homolog pmGal.2 - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 07-Dec-1999
C:Accession: A49218
R:Markham, P.F.; Glew, M.D.; Whithear, K.G.; Walker, I.D.
A:Title: Molecular cloning of a member of the gene family that encodes pmGA, a hemagglut
A:Reference number: A49218; MUID:93162830
A:Accession: A49218
A:Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 1-647 <MAR>
A:Cross-references: GB:S55216; NID:g265625; PIDN:AAB25397.1; PID:g265626
A:Experimental source: S6
A:Note: sequence extracted from NCBI backbone (NCBIN:125182, NCBI:P:125183)
C:Genetics:
A:Genetic code: SGC3

Query Match 35.7%; Score 713; DB 2; Length 647;
Best Local Similarity 41.1%; Pred. No. 8.8e-32;
Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps 12;

QY 8 DANPNNGO-----TQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTN 60
DB 48 DTNPGDGGGMMNAASQELAAARMGLTTFDSKAKNLGLYVDYKKTQNTLTTRAYDAAKTVL 107
QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVAYKALKTTLEQRAATNLEGLS 120
DB 108 DNSSSTTQNLNEAKTRLETAIRTAATSQTDFEQHAELVKVYKELKTTLSNETATLAPYA 167
QY 121 STAYNQIRNNLDVLYNKASSLTITKLDPLNGTGLDLSNEITTANKNINNTL--STINQOK 178
DB 168 DAQYAGIKMHLISGLYDAGKAITTKLEPVEGDP--LTASAVMANTKIIEAIDKDEVLPQK 226
QY 179 TNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPV-----NKKY 230
DB 227 ENATKLADSFVKQVLVKEKITGVEEAAH--KAQPNYSFVGISVDITGTTGTGTSIPNWDY 285
QY 231 ARRTVW--NGDEPSSRILANT-----NSITDVSWIYSLAGTNTKYQFSFNSYGPSTGY 281
DB 286 AORTIFTNSDEP--RSISNTPADGOTMAQPLSNVSWIYSLAGTGAKYKTLFTFYGPSTGY 343
QY 282 LYFPYKLVKAADANNVGLQYKLNNGNVQOEFATSTSA-----NNTTANPTPAVDEIKVA 336
DB 344 LYFPYKLVNTSDQVKGLEYKLD-----ATKPSAITFGSDQTMNGKTPTVNDINVA 395
QY 337 KIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIGNYSNNENNADKI 383
DB 396 KVTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 437
RESULT 5
A:hemagglutinin homolog pmGal.2 - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 07-Dec-1999
C:Accession: A49218
R:Markham, P.F.; Glew, M.D.; Whithear, K.G.; Walker, I.D.
A:Title: Molecular cloning of a member of the gene family that encodes pmGA, a hemagglut
A:Reference number: A49218; MUID:93162830
A:Accession: A49218
A:Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 1-647 <MAR>
A:Cross-references: GB:S55216; NID:g265625; PIDN:AAB25397.1; PID:g265626
A:Experimental source: S6
A:Note: sequence extracted from NCBI backbone (NCBIN:125182, NCBI:P:125183)
C:Genetics:
A:Genetic code: SGC3

QY 278 STGYLYFPYKLVKAADANNVGLQYKLNNGNVQOEFATSTSA-----NNTTANPTPAVDE 332
DB 338 STGYLYFPYKLVNTSDQVKGLEYKLD-----ATKPSAITFGSDQTMNGKTPTVND 389
QY 333 IKVAKIVISGLRFGQNTIELSVPTGEGNNKVPAMIGNIGNYSNNENNADKI 383
DB 390 INVAKVTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 435
RESULT 6
S48755
Major surface protein (clone pmGal.5) precursor - Mycoplasma gallisepticum (fragment)
C:Species: Mycoplasma gallisepticum
C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S48755
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.
A:Title: The organisation of the multigene family which encodes the major cell surface
A:Reference number: S48751; MUID:95010739
A:Accession: S48755
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-386 <MAR>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62419.1; PID:g535692
A:Note: the sequence of residues 385-386 and the corresponding nucleotide sequence ar
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match 28.8%; Score 575.5; DB 2; Length 386;
Best Local Similarity 42.3%; Pred. No. 1.5e-24;
Matches 138; Conservative 50; Mismatches 113; Indels 25; Gaps 10;

QY 8 DANPNNG-----TQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEAEVTNN 61
DB 64 NTNPGCGGMDNSAQQLAAAKKELSDLLATQNSNLSTYADYAKIKNDLTAAYTTAETASQ 123
QY 62 NUNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVAYKALKTTLEQRAATNLEGLS 121
DB 124 NQAAATLEQVKNAASTLQTAINTAVNEKVFEDENNSLVTAYTNLKTLEGEENTTAAFN 183
QY 122 TA-YNOIRNNLDVLYNKASSLTITKLDPLN--GTLTLDLSNEITTANKNINNTL--STINEQ 177
DB 184 SANYGKIKTHLSLVNQAKTITTSIL--LNDAGQSPKNDVNVKINKEITDAINPTLLNQ 241
QY 178 KTNADALSNSFIKKVIONNEQSFVG---TFTNANVOPSNYSFVAFSADVTP-----VNY 228
DB 242 KANADWLATSFTKQVL--NDAQLTSGSSETSMQTQPGNYSFVGYSVDVTTGSSNNARPNN 300
QY 229 KYARTVWNGDEPSSRILANTNSITDVSWIYSLAGTNTKYQFSFNSYGPS--TCYLFPYK 287
DB 301 NFAQRKVDNTNRAPLAQTEQSKLTDVSWIYSLSGMGAKYTTVTFDYIGASNNAYLFPYK 360
QY 288 LVKAADANNVGLQYKLNNGNVQOEF 313
DB 361 LVQTNQ--NVGLQYLVNNTTPKLVNF 384
RESULT 7
S51560
Major surface protein (clone pmGal.6) - Mycoplasma gallisepticum (fragment)
C:Species: Mycoplasma gallisepticum
C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S51560; S48757
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.
A:Title: The organisation of the multigene family which encodes the major cell surface
A:Reference number: S48751; MUID:95010739
A:Accession: S51560
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA

1

1

Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2422
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-661 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA076634.1; PID:g17134073; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all4935

Query Match 8.3%; Score 165.5; DB 2; Length 661;
Best Local Similarity 22.7%; Pred. No. 0.082;
Matches 88; Conservative 40; Mismatches 188; Indels 71; Gaps 11;
Qy 18 LEAARMELTDLI-----NAKAMTLASLDYAKIEASLSAYSEAFVNNNNLNATLEQ 69
Db LKKAEDDLSGIIRSTNRAKSDANSRLNRAEQDETFQADAGFOTATLNTATNNRNNQA 123
Qy 70 LKMAKTNLESINQANTDKTFDEHEPNLYEAYKALKTLEQATNLEGLSSTAYNQIRN 129
Db 124 LNTATNNRNNQAALNTATNNRDN-----AQNALNTATNNR--NNAQNALNTATNN--RN 174
Qy 130 NLVDLYNKASSLITKTLDPGLNGTLLDSNEITTKANKNNITLSTINEQKTNADALSNSFI 189
Db 175 NAQNALNTATNNRNNQAAL-----NTATNNRNNQAALNTATNNRNNQAALNTATNN 226
Qy 190 KKVQI-----NNRQSPVGTFTNANVQPSYSFVAFSADVTPVNYKYARRTVWNGDEPSSRIL 246
Db 227 RNLQAARRGNSQKAIQAINALNQALNTATNTALNTATNTANTAFN-----T 278
Qy 247 ANTSITDVSIIYSLAGTNTKQFSFSGYSGYFPPYKLVKAADANNVGLQYKLNNG 306
Db 279 ANNFFST-----ANTAFNTATNNFSTA-----NTALNTATNNF 311
Qy 307 NVQOVEATSPSANTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVP 366
Db 312 NTAAELDQANRLTARN-----DFTANSFRTGNELTATNTATNNFNTANTFTATT 365
Qy 367 MIGNIYSSNNENNADKIPGYRRPGTFL 393
Db 366 NFNN--ASSRNTAEQARNQVRETRL 390

RESULT 11
D90011
FmtB protein [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D90011
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2481 <KUR>
A:Cross-references: GB:BA000018; PID:g13701961; PIDN:BA043253.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fmcB(mrp)

Query Match 8.2%; Score 164.5; DB 2; Length 2481;
Best Local Similarity 23.1%; Pred. No. 0.54;
Matches 98; Conservative 56; Mismatches 156; Indels 115; Gaps 20;

Qy 8 DANPNNGQTOLE-----AARMELTDLINAKAMTLASLDYAKIEASLSAYSEAE 57
Db 1186 DQATTGETSIDQVTPPTVKKKATARNETAILNNKLEIQATPDATDEKQAADA--EAN 1243
Qy 58 TVNNNLN-----ATLEQLKMAKTNLESAINQANTDKTTFDEHEPNLYEAYKALKTTL-- 109
Db 1244 TENKANQALSAATTAQAQVDEAKANAEAAIN-AVTPKYVKKQAADKDEIDQLQATQTNVIN 1302
Qy 110 -EQATNLEGLS-----STAYNQIRNL-----VDLYNKASSLITKTLDPGLNGTLL 155
Db 1303 NQATNLEEKAAIQQLATAVTDAKNNITATDONGVDATKAGKNSIQSTQP---ATAV 1359
Qy 156 DSN-----BITTANKNNITLSTINEQKTNADALSNSFIKKVIONNEQSPVGTFTNAN 208
Db 1360 KSNNAKNEVDQAVTQNAIDNTGATTEKNAAKDL-----VLKAKAKAYODIL--N 1409
Qy 209 VQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRIILANTNSITDVSIIYSLA----GT 264
Db 1410 AQTNN-----DVTQI-----KQAVADIQGITADTTIKDVAKDELAT 1446
Qy 265 NTKYQFSFSGYSGYFPPYKLVKAADAN-----NVGLQYKLVNGNVQOVEFATST 317
Db 1447 KANEQKAL-----TAQTADATTEKEQANQVDAQLTOGN-QNIEAQSII 1490
Qy 318 SANNTTA-NPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSN 376
Db 1491 DDVTAKDNAIQADPIQASTDVKTNR-----ABELTEM-----QNKITEILNNNETTNE 1541
Qy 377 ENNAD 381
Db 1542 EKGND 1546

RESULT 12

C89921
hypothetical protein ebhB [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89921
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C89921
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3890 <KUR>
A:Cross-references: GB:BA000018; PID:g13701233; PIDN:BA042528.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebhB

Query Match 8.2%; Score 163.5; DB 2; Length 3890;
Best Local Similarity 24.4%; Pred. No. 1.1;
Matches 94; Conservative 54; Mismatches 153; Indels 85; Gaps 20;

Qy 16 TQLEAARMELTDLINAKAMTLASLDYAKIEASLSAYSEA-ETVNNNLNATLEQLKMAK 74
Db 3528 TDLNTAMGNLQGAINDTEQTLNS-QNYQDATPSKKTATYNAVQAADKILNKSQGNK-TK 3585
Qy 75 NLESAINQANTDKTTFDEHEPNLYEAYKALKTLEQATNLEGLSSTAYNQIRNLVDL 134
Db 3586 DQVTEAMNQVNSAKNNLDG-----TRLLDQAKQAKQQLNNMTHTLTAAQKTLNQ---- 3636
Qy 135 YNKASSLITKTLDPGLNGTLLDSNEITTKANKN-INNTLSTINEQKTNADA--LSNSFI-- 189
Db 3637 -----INSGTTVAGVHTVOSNANTLQDAMNTLRQSTIANNDATKASEDYDA 3682
Qy 190 ---KKVIONNEQSPVGTFTNANVQPS-NYSFVAFSADVTPVNYKYARRTVWNGDE----- 240

Db 430 ESNVSDQTSSEAS-TNSNSISLSPSNISSTSD--SESATNSSDFSNVAEVANNSLASV 486
Qy 64 -NATLEQLKMAKTLESAINQANTDKTFDNEHPNPLVEAYKALKTTLEQRATN----- 115
Db 487 NNSSSVLSSTSTADNLGINSQSDNLTQDSEISTSGAFLSSNOTSEASTNSSSISL 546
Qy 116 -----LEG-LSSTAYNOI-----RNNLDVLYNKASSLI--TKTLDPL-----NGG 152
Db 547 SPSNISSTSVLESTTSSSNFENVAEVANNSLASVNNSSSVLSSTSTADNLEINQFGSDN 606
Qy 153 TLLDSNEITTANKNINNTLSTINEQKTNADALSNSFIKKVIONNEQSPVGTFTNANVOPS 212
Db 607 LTKDSSEISTSG-----AFLSSNOTSEASSNS-----MSSINSPSLSLTNSSESAT 655
Qy 213 NVSFVAFSADYTPVNYKYARTVWNGDEPSSRIL-----ANTNSITDVSMIYSLAG 263
Db 656 NQS-----NSSEATKVDN-----NSSTHSSNII NSGNSDSDSDSDSSNLSLSPN 704
Qy 264 TNTKYQFSFNYGPGTGYLPFYKLVKAADANNVGLQYKLNNGNVQOVEFATSTSANNT- 322
Db 705 LETNOTIS-----SKPSEVNNI-----SENPKVSSNSVQENSTD 740
Qy 323 ---TANPTPAVDE-----TKVAKIVLSGLRFGQNTTE---LSVPTGEG 359
Db 741 HEMSTNPKRSISSTSTSSQKESQSNLLNTTTEGINNPITFNNSSENSESAASILTYS 800
Qy 360 NMKVAPMIGNIYLS-----SNENNADKI 383
Db 801 NNSSESETGCLYISNEAQRONGSEISHLPSSNSENNVSSI 843

Search completed: June 12, 2002, 10:49:47
Job time: 205 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:45:47 ; Search time 22.56 Seconds
(without alignments)
425.499 Million cell updates/sec

Title: US-09-147-052-2_COPY_64_456

Perfect score: 1998

Sequence: 1 CMSITKDKANNGQTLEA.....SSNENADKIPCYRRPGTFL 393

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1910	95.6	615	2	US-08-525-742-10
2	1856	92.9	610	2	US-08-525-742-8
3	1638	82.0	368	2	US-08-525-742-4
4	1612	80.7	368	1	US-08-185-851A-4
5	806	40.3	661	2	US-08-525-742-2
6	159	8.0	933	3	US-08-293-728-2
7	159	8.0	933	4	US-09-421-868-2
8	136	6.8	1098	4	US-08-923-992A-8
9	135.5	6.8	1073	4	US-09-541-782-6
10	132.5	6.6	1002	4	US-09-268-347-24
11	131.5	6.6	1164	4	US-08-923-992A-2
12	131	6.6	1004	4	US-09-268-347-30
13	131	6.6	1104	4	US-08-923-992A-4
14	129	6.5	2048	4	US-09-268-347-48
15	126.5	6.3	1128	4	US-08-923-992A-6
16	126.5	6.3	1164	4	US-08-923-992A-10
17	125.5	6.3	518	3	US-09-043-123-2
18	125.5	6.3	3788	4	US-09-336-447A-76
19	124.5	6.2	2411	4	US-09-268-347-36
20	122.5	6.1	1536	1	US-08-038-682-2
21	122.5	6.1	1536	1	US-08-302-832-2
22	122.5	6.1	1536	2	US-08-530-198-2
23	122.5	6.1	1536	2	US-08-469-880-2
24	122.5	6.1	1536	2	US-08-728-470-2
25	122.5	6.1	1536	2	US-08-617-697-2
26	122.5	6.1	1536	4	US-08-719-641-2
27	122.5	6.1	1565	6	5352450-2

28	120.5	6.0	746	5	PCT-US95-10509-2
29	120.5	6.0	1104	4	US-09-268-347-28
30	120.5	6.0	1104	4	US-09-268-347-34
31	119.5	6.0	984	1	US-08-242-932-2
32	119.5	6.0	984	1	US-08-714-481-2
33	119.5	6.0	984	5	PCT-US95-06111-2
34	119.5	6.0	2314	4	US-09-268-347-49
35	119	6.0	2354	4	US-09-268-347-47
36	118.5	5.9	866	2	US-08-483-101-4
37	118	5.9	1612	1	US-08-169-927-2
38	117.5	5.9	889	4	US-09-336-447A-15
39	117.5	5.9	1912	1	US-08-409-995-4
40	117.5	5.9	1912	3	US-08-685-467-4
41	117	5.9	2353	4	US-09-377-155-33
42	117	5.9	2353	4	US-08-913-942-4
43	117	5.9	2353	4	US-09-669-974-33
44	116	5.8	941	4	US-09-336-447A-9
45	116	5.8	1338	2	US-08-728-470-9

ALIGNMENTS

RESULT 1

US-08-525-742-10

; Sequence 10, Application US/08525742

; Patent No. 5871742

; GENERAL INFORMATION:

; APPLICANT: Saito, Shuji

; APPLICANT: Ohkawa, Setsuko

; APPLICANT: Saeki, Sakiko

; APPLICANT: Ohsawa, Ikuroh

; APPLICANT: Funato, Hiroo

; APPLICANT: Iritani, Yoshikazu

; APPLICANT: Aoyama, Shigemi

; APPLICANT: Takahashi, Kiyoohito

; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE

; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND

; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL

; TITLE OF INVENTION: AS USE THEREOF

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &

; STREET: 1725 K Street, Suite 1000

; CITY: Washington

; STATE: DC

; ZIP: 20006

; COUNTRY: USA

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/525,742

; FILING DATE: 25-SEP-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 05-074139

; FILING DATE: 31-MAR-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 05-245625

; FILING DATE: 30-SEP-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/JP94/00541

; FILING DATE: 31-MAR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Mclelland, Le-Nhung

; REGISTRATION NUMBER: 31,541

; REFERENCE/DOCKET NUMBER: 950811

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-659-2930

```

; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-525-742-10

Query Match          95.6%; Score 1910; DB 2; Length 615;
Best Local Similarity 98.2%; Pred. No. 3.8e-132;
Matches 376; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 60
Db 27 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 86
QY 61 NNLNATLEQLKAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
Db 87 NNLNATLEQLKAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
QY 121 STAYNOIRNNLDVLYNKASSLITKTLDPNGGTLSDSNEITTKANKNINNTLSTINEQKTN 180
Db 147 STAYNOIRNNLDVLYNKASSLITKTLDPNGGMLSDSNEITTVNRNINNTLSTINEQKTN 206
QY 181 ADALNSFIKKVIQNNNEQSFVGTFTNANQVPSNYSFVAFSADVTPVNYKYARTVWNGDE 240
Db 207 ADALNSFIKKVIQNNNEQSFVGTFTNANQVPSNYSFVAFSADVTPVNYKYARTVWNGDE 266
QY 241 PSSRIILANTNSITDVSIIYSLAGTNTKYQFSFSNYGSPSTGYLYFPYKLYKAADANNVGLQ 300
Db 267 PSSRIILANTNSITDVSIIYSLAGTNTKYQFSFSNYGSPSTGYLYFPYKLYKAADANNVGLQ 326
QY 301 YKLNNNGVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 360
Db 327 YKLNNNGVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 386
QY 361 MNKVAPMIGNIYLSNENNAKDI 383
Db 387 MNKVAPMIGNIYLSNENNAKDI 409

RESULT 2
US-08-525-742-8
; Sequence 8, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hirono
; APPLICANT: Iritani, Yoshikazu
; APPLICANT: Aoyama, Shigemi
; APPLICANT: Takahashi, Kiyohito
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
; TITLE OF INVENTION: AS USE THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
; ADDRESSEE: NAUGHTON
; STREET: 1725 K Street, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,742
; FILING DATE: 25-SEP-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-074139
; FILING DATE: 31-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 05-245625
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/00541
; FILING DATE: 31-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 950811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-659-2930
; TELEFAX: 202-8870357
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-525-742-8

Query Match          92.9%; Score 1856; DB 2; Length 610;
Best Local Similarity 95.5%; Pred. No. 3.4e-128;
Matches 365; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

QY 1 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 60
Db 27 CMSITKDDANPNNGQTQLEAARMELTDLINAKAMTILASLDYAKIEASLSAYSEAEVTN 86
QY 61 NNLNATLEQLKAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
Db 87 NNLNATLEQLKAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 146
QY 121 STAYNOIRNNLDVLYNKASSLITKTLDPNGGTLSDSNEITTKANKNINNTLSTINEQKTN 180
Db 147 STAYNOIRNNLDVLYNKASSLITKTLDPNGGTLSDSNEITTKANKNINNTLSTINEQKTN 206
QY 181 ADALNSFIKKVIQNNNEQSFVGTFTNANQVPSNYSFVAFSADVTPVNYKYARTVWNGDE 240
Db 207 ADALNSFIKKVIQNNNEQSFVGTFTNANQVPSNYSFVAFSADVTPVNYKYARTVWNGDE 266
QY 241 PSSRIILANTNSITDVSIIYSLAGTNTKYQFSFSNYGSPSTGYLYFPYKLYKAADANNVGLQ 300
Db 267 PSSRIILANTNSITDVSIIYSLAGTNTKYQFSFSNYGSPSTGYLYFPYKLYKAADANNVGLQ 326
QY 301 YKLNNNGVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 360
Db 327 YKLNNNGVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 386
QY 361 MNKVAPMIGNIYLSNENNAKDI 382
Db 387 MNKVAPMIGNIYLSNENNAKDI 408

RESULT 3
US-08-525-742-4
; Sequence 4, Application US/08525742
; Patent No. 5871742
; GENERAL INFORMATION:
; APPLICANT: Saito, Shuji
; APPLICANT: Ohkawa, Setsuko
; APPLICANT: Saeki, Sakiko
; APPLICANT: Ohsawa, Ikuroh
; APPLICANT: Funato, Hirono

```


APPLICANT: Iritani, Yoshikazu
 APPLICANT: Aoyama, Shigemi
 APPLICANT: Takahashi, Kiyoohito
 TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
 TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
 TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
 TITLE OF INVENTION: AS USE THEREOF
 NUMBER OF SEQUENCES: 51
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
 ADDRESSEE: NAUGHTON
 STREET: 1725 K Street, Suite 1000
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/525,742
 FILING DATE: 25-SEP-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 05-074139
 FILING DATE: 31-MAR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 05-245625
 FILING DATE: 30-SEP-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/JP94/00541
 FILING DATE: 31-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Mclelland, Le-Nhung
 REGISTRATION NUMBER: 31,541
 REFERENCE/DOCKET NUMBER: 950811
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-659-2930
 TELEFAX: 202-8870357
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 368 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-525-742-4

Query Match 82.0%; Score 1638; DB 2; Length 368;
 Best Local Similarity 95.5%; Pred. No. 1.5e-112;
 Matches 322; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
 Qy 1 CWSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 60
 Db 27 CWSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 86
 Qy 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
 Db 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
 Qy 121 STAYNQIRNNLVLYNKASSLITKTLDPNGGTLSDSNEITTANKNINNTLSTINEQKTN 180
 Db 147 STAYNQIRNNLVLYNNASSLITKTLDPNGGMLSDSNEITTVNRNINNTLSTINEQKTN 206
 Qy 181 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSDVTPVNYKYARRTVWNGDE 240
 Db 207 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSDVTPVNYKYARRTVWNGDE 266
 Qy 241 PSSRLANTNSITDVSWIYSLAGTNTKQFSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 300
 Db 267 PSSRLANTNSITDVSWIYSLAGTNTKQFSFNSYGPSTGYLYFPYKLVKAADANNVGLQ 326

Qy 301 YKLNNNGVQVEFATSTSANNTTANPTPAVDKIKVAK 337
 Db 327 YKLNNNGVQVEFATSTSANNTTANPTQOLMRKLK 363
 RESULT 4
 US-08-185-851A-4
 Sequence 4, Application US/08185851A
 Patent No. 5489430
 GENERAL INFORMATION:
 APPLICANT: Saito, Shuji
 APPLICANT: Ohkawa, Setsuko
 APPLICANT: Fujisawa, Ayumi
 APPLICANT: Iritani, Yoshikazu
 APPLICANT: Aoyama, Shigemi
 TITLE OF INVENTION: Poultry Mycoplasma Antigens, Gene
 TITLE OF INVENTION: Thereof and Recombinant Vectors Containing the Gene As
 TITLE OF INVENTION: Well As Vaccines Utilizing the Same
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &
 ADDRESSEE: Naughton
 STREET: 1725 K Street, N.W., Suite 1000
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS version 5.0
 SOFTWARE: ASCII from Word Perfect version 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/185.851A
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Theresa M. Stevens-Smith
 REGISTRATION NUMBER: 36,281
 REFERENCE/DOCKET NUMBER: PO-8-A930918
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-659-2930
 TELEFAX: 202-887-0357
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 368 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-185-851A-4

Query Match 80.7%; Score 1612; DB 1; Length 368;
 Best Local Similarity 95.0%; Pred. No. 1.2e-110;
 Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
 Qy 1 CWSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 60
 Db 27 CWSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 86
 Qy 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 120
 Db 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
 Qy 121 STAYNQIRNNLVLYNKASSLITKTLDPNGGTLSDSNEITTANKNINNTLSTINEQKTN 180
 Db 147 STAYNQIRNNLVLYNNASSLITKTLDPNGGMLSDSNEITTVNRNINNTLSTINEQKTN 206
 Qy 181 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSDVTPVNYKYARRTVWNGDE 240
 Db 207 ADALNSFIKKVQIONNEQSFVGTFTNANVQPSNYSFVAFSDVTPVNYKYARRTVXNGDE 266

QY 241 PSSRILANTNSIDYSWIYSLAGTNTKQFSPSNYGPSTGYLYFPYKLVKAADANNVGLQ 300
 Db 267 PSSRILANTNSIDYSWIYSLAGTNTKQFSPSNYGPSTGYLYFPYKLVKAADANNVGLQ 326
 QY 301 YKLNNGNQVQVEFATSTSANNTTNPATPAVDEIKVAK 337
 Db 327 YKLNNGNQVQVEFATSTSANNTTNPATPAVDEIKVAK 363

RESULT 5
 US-08-525-742-2
 ; Sequence 2, Application US/08525742
 ; Patent No. 5871742
 ; GENERAL INFORMATION:
 ; APPLICANT: Saito, Shuji
 ; APPLICANT: Ohkawa, Setsuko
 ; APPLICANT: Saeki, Sakiko
 ; APPLICANT: Ohsawa, Ikuroh
 ; APPLICANT: Funato, Hirono
 ; APPLICANT: Iritani, Yoshikazu
 ; APPLICANT: Aoyama, Shigemitsu
 ; APPLICANT: Takahashi, Kiyoo
 ; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
 ; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
 ; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
 ; TITLE OF INVENTION: AS USE THEREOF
 ; NUMBER OF SEQUENCES: 51
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
 ; ADDRESSEE: NAUGHTON
 ; STREET: 1725 K Street, Suite 1000
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/525,742
 FILING DATE: 25-SEP-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 05-074139
 FILING DATE: 31-MAR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 05-245625
 FILING DATE: 30-SEP-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/JP94/00541
 FILING DATE: 31-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Mclelland, Le-Nhung
 REGISTRATION NUMBER: 31,541
 REFERENCE/DOCKET NUMBER: 950811
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-659-2930
 TELEFAX: 202-8870357
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 661 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

Query Match 40.3%; Score 806; DB 2; Length 661;
 Best Local Similarity 42.8%; Pred. No. 2,8e-51;
 Matches 185; Conservative 62; Mismatches 123; Indels 62; Gaps 12;

QY 8 DANPNNGQTQLEAARMELTD-----LINAKMTLASQDYAKTEASLSAYSEAEV 60
 Db 45 NTFNSDGGMMNAARAKELADAKAALTTILINGETANLASDYAKIKSELTAYETAKAVS 104
 QY 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLS 120
 Db 105 AKTGATLNEVNEAKTTLDAALKKAASAKNDFDAQHGLSVEAYNNLKETLKEEKNLDSLA 164
 QY 121 STAYNOIRNNLVLYNKASSLITKTLPLNGGTLDSNEITTTANKNINNTTNEOKTN 180
 Db 165 NENYAAIRTNLSLYEKANTIVTATLDPAT-GNIPVMSVTQANQDITNATSRILAKWN 223
 QY 181 ADALNSFIKKVIONNEQSFYGTFTNANVQPSNTSFVAFSADVTPVNVKYARRVWNGDE 240
 Db 224 ADNLANSFIKQSLVKNLIRYDVANNOE-QPANYSFVGFVSVNVDTPNWNFAQRKVAWSEN 282
 QY 241 -----PSSRILANTNSIDYSWIYSLAGTNTKQFSPSNYGPSTGYLYFPYKLVKAAD 293
 Db 283 TPLATTPAEDATQAAALTDVSWIYSLAGAEAKYTLTSFRVGAETAYLYFPYKLVKTS 342
 QY 294 ANNVGLOYKLNGNVQVEF-----ATSTSAN-NTTANP-----TPAVDEIKVAKIV 339
 Db 343 --NVGLOYKLNGGDKTKOINFTVQTPASGSDVAANEETMASPAEQMSAPTVDIKIARVA 400
 QY 340 LSGURFGONTIELSVPTGEGNMKNKVPIMIGNIYLS-----NENN 379
 Db 401 LSNLKFSNTIEFSVPTG-----KAAPMIGNMYLTSSSEVKNKIYDDLFGNSFNENN 455
 QY 380 ----ADKIPY 386
 Db 456 PTAVTVLLKGY 467

RESULT 6
 US-08-293-728-2
 ; Sequence 2, Application US/08293728D
 ; Patent No. 6008341
 ; GENERAL INFORMATION:
 ; APPLICANT: Foster, Timothy J.
 ; APPLICANT: McDevitt, Damien L.
 ; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
 ; FILE REFERENCE: 05344.105011
 ; CURRENT APPLICATION NUMBER: US/08/293,728D
 ; CURRENT FILING DATE: 1994-08-22
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 933
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 US-08-293-728-2

Query Match 8.0%; Score 159; DB 3; Length 933;
 Best Local Similarity 22.8%; Pred. No. 0.001;
 Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;

QY 2 MSITKKDANPNNGQTQL--EAARMELTDLINAKMTLASQDYAKTEASLSAYSEAEV 59
 Db 73 VSDTKSSNTNNGETSVQNPQAQETTQSSSTNATI-----EETPTVGEATT 120
 QY 60 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGL 119
 Db 121 TTNOANTPATTOSSNTNABELVNO-TSNETTEND--TNTVSSVNS-----PQNSTNAENV 172
 QY 120 SST-----AYNOIRNNLVLYNK-----ASSLITKTLPLNGGTL 155
 Db 173 SITQDPTSTATPSNNEAPQSDASKNDVNVQAVNTSAPMRAFSLAAVAADAPAGTDI 232
 QY 156 DSNEITTTANKNINNTLSTINEQ-----KTN-ADALNSFIK-----KVIONNEQSFVGTFTN 206
 Db 233 -TNLTNVTGIDSGTIVYVPHQAGYVKNYGFVSPNSAVKGDFTKITVPKELNNGVTST 291

QY 207 ANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRLANTNSITDVSIIYSLAG-TN 265
Db 292 AKVPP-----IMAGDQ-----VLANGVIDSGNVIYTFDYN 324
QY 266 TKYQFSFNSGPGCYLYFPYKLVKAADANNVGLQYKLNNGNVQVQEFATSTSANNTAN 325
Db 325 TKDDVKATLTMPA--YI-----DPENV-----KKTGNV-----TLATGIGSTTAN 362
QY 326 PTPAVDEIKVAKI-----VLSGLRFGQNTIELSVPTGEGNMKNKVPAMI-GNIVLSNE 377
Db 363 KTVLVDEYKYGKFNLSIKGTIDQIDKTNNTYRQTIYVNPSSGDNVIAVLGNLKPNTDS 422
QY 378 N 378
Db 423 N 423
RESULT 7
US-09-421-868-2
; Sequence 2, Application US/09421868
; Patent No. 6177084
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/09/421.868
; CURRENT FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 08/293,728
; PRIOR FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-421-868-2
Query Match 8.0%; Score 159; DB 4; Length 933;
Best Local Similarity 22.8%; Pred. No. 0.001;
Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;
QY 2 MSITKKDANPNNGQTL--EAARMELTDLINAKAMTLASLDYAKIEASLSAYSEATV 59
Db 73 VSDTKTSNTNGETSVANPAQQTSSNTATT-----EETPVTCETTT 120
QY 60 NNNLNLATLEQLKMAKTNLESAINQANTDKTTFDNEHPLNVEAYKALKTTLEQRTNLEGL 119
Db 121 TTNOANTPATQSSNTAEELVNO-TSNETTFND--TWTSSVNS-----PQNSINAENV 172
QY 120 SST-----AYQIRNRLVDLYNK-----ASSLTKTLDPLNGGTL 155
Db 173 STTQDTSTPATPSNNESAPQSDASKNQVAVNTSPRMAFSLAAVAADAPAGTDI 232
QY 156 DSNETTTANKNNITLSTINEQ-----KTN-ADALNSFIK---KVIQNNQSFVGTFTN 206
Db 233 -TNQLTNTVIGDSGTTVYPHQAGVVKLNLYGFSVPNSAVKGDTFKITVPKELNMGVST 291
QY 207 ANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRLANTNSITDVSIIYSLAG-TN 265
Db 292 AKVPP-----IMAGDQ-----VLANGVIDSGNVIYTFDYN 324
QY 266 TKYQFSFNSGPGCYLYFPYKLVKAADANNVGLQYKLNNGNVQVQEFATSTSANNTAN 325
Db 325 TKDDVKATLTMPA--YI-----DPENV-----KKTGNV-----TLATGIGSTTAN 362
QY 326 PTPAVDEIKVAKI-----VLSGLRFGQNTIELSVPTGEGNMKNKVPAMI-GNIVLSNE 377
Db 363 KTVLVDEYKYGKFNLSIKGTIDQIDKTNNTYRQTIYVNPSSGDNVIAVLGNLKPNTDS 422
QY 378 N 378

Db 423 N 423
RESULT 8
US-08-923-992A-8
; Sequence 8, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923.992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-8
Query Match 6.8%; Score 136; DB 4; Length 1098;
Best Local Similarity 22.7%; Pred. No. 0.063;
Matches 84; Conservative 51; Mismatches 159; Indels 76; Gaps 14;
QY 4 ITTKDANPNNGQTL--EAARMELTDLINAKAMTLASLDYAKIEASLSAYSE-----AE 57
Db 192 IRKQAQPDKKEDAEVKEELGKFSSTKAGLDQIEQHVKKETSSSEENTQKVDEHYAN 251
QY 58 TVNNNLNATLQLKMAKTNLESAINQANTDKTTFDNEHPLN-----VEAYKALK 106
Db 252 SLQNLQAQKSLLELDKATTN-----EQATQVKNOFLNAQKLEIQIPLIKETNVKLYKAMS 306
QY 107 TLEQRTNLEGLSTAYNQIRNNRLVDLYNKASSLITKTTLDPLNGTLDLSNEITANKN 166
Db 307 ESLEQVEKELKHNSEA-----NLEDLVAKSEIYREYEGKLNQKNLP--ELQKLEEE 357
QY 167 INNTLSTINEQ-----KTNADALNSFIKKVI---QNNQSFVGTFTNANVQPSNYSFVA 218
Db 358 AHSKLKQVVEDFRKKFKTSEQVTPKKRLKRDLAANNQKKI-----ELTVSPENI---- 408
QY 219 PSADVTVPYKYARTVWNGDEPSSRLANTNSITDVSIIYSLAGTNTKYQFSFS----- 273
Db 409 -----TVYEGEDVKFTVAKSDKTTLDFDILL-----TKYNPVSVDRIST 449
QY 274 NYGPSTG---YLYFPYKLVKAADANNVGLQYKLNNGNVQVQEFATSTSANNTANP-TPA 329

```

; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-24

Query Match
Best Local Similarity 6.6%; Score 132.5; DB 4; Length 1002;
Matches 91; Conservative 58; Mismatches 153; Indels 175; Gaps 21;

QY 13 NGOTOLEAARM---ELTDLINAKAMTLASLODYAKIEASLSAYSE-----AST 58
Db 362 NGTNPVKISNVADGTTDAVSEK---QLKALQD-KQVTLASNAYANGSDADGKATQT 418
QY 59 VNNLNATLEOLKMAKTNLESA-----INQANTDKTTFDN-----EHPN 97
Db 419 LGNDLNFKEKSTDELLNIKAAGDTVTTPKKGSVQVGDGKATIQDGAKTITGLVEASE 478
QY 98 LYEAYKAL-----KTTL-----EQATN----- 115
Db 479 LVDSLNLKLGKVGKDGTDGTHDTLVKSGDKVTLKAGDNLKVKQEGTFTYVLRD 538
QY 116 -LEGSSSTAYNQIRNNLVLYNKASSLITK---TLDP---LNGGTLDSNETTANKNIN 168
Db 539 ELTGKVSVEFKTENGCA-----NGASTKIYKDGTLTTPANDANGAATDADKIKVASDGI- 593
QY 169 NTLSTINEOKTNADLSNFIKKVIONNEQSFVGTFTNANQVPSNYSFVAFSADVTPVNY 228
Db 594 -----SAGNKAVKNV-----SGLKKFGDANFN-----LTSSADNLTKOY 629
QY 229 KYARRTVWNGDEPS-----SRILANTNSITDVSWIYS---LAGTWTKYQSFESNYG 276
Db 630 DNAYKGLTLDERSKQKQPTVADNTAATVGDRLGWLVSADKTTGESKEYSAQVRNAN 689
QY 277 PSTGYLYFPYKLVKAADANNVGLQYKLNNGNQVQVEFATSTANNTANTPTPAVDEIKVA 336
Db 690 E-----VKFKSGNGINVSCKTLONGTRETIFELAKDEN----- 722
QY 337 KIVLSGLRFGQNTIEL---SVPTGECNM-----NKVAPMIGNIYLSSNEN 378
Db 723 -----AIAFGSGSKALRDNTVAIGTGNVNVNAEKSAGFDPNTIEDKAGSYAFGNDN 774

RESULT 11
US-08-923-992A-2
; Sequence 2, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-Iga Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
```

```

; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-24

Query Match
Best Local Similarity 6.8%; Score 135.5; DB 4; Length 1073;
Matches 97; Conservative 76; Mismatches 182; Indels 109; Gaps 21;

QY 13 NGO-TOLEAARMELTDLINAKA-MTLASL---QDYAKIEASLSAYSEATVNNLNAT 66
Db 611 NGYFTLLDNFNASWELLNTHSNQLLISMTKITEHFQSLIDEALQSARSSCAVENSLLDI 670
QY 67 LEOLKMAKTNLESAINQANTDKTTFDNEHPN-----LVEAYKALKIT---LEORATNLEG 118
Db 671 VSELKSKNSLLDALEHSLQDTSMSOKLNGISSELIELQDKMESYROLVQELRSLYN 730
QY 119 LSSTAYNQIRNNLVLYNKASSLITKTLPLNGGTLDSNEIP-----TA 163
Db 731 LQHTHEESQKELMGVNRNDIDALVKTCTSLNDADIILSDYISDQSKFESKQODLIANI 790
QY 164 NKINNTLSTINEQ-KTNADAL-----SNSFIKK---VIONNEQSFVGTFTNANQVPSN 213
Db 791 GKIVSNFLOQNESLYTKADILHSLNDTNSIRKANEIMNRSSEFL---RNA----- 841
QY 214 YSFVAFSADVTPVNYKYARTVWNG-----DEPSSRILANTNSITDVSWIYSLA----- 262
Db 842 -----ASQAEIVGANKERIOKTVENGSQLDLSKSKATHNSRMYD---HCLALAESQKQ 893
QY 263 GTWTKYQ-----PFSWYGPSTGYLYPPYKLVKAADANNVGL---QYKLNNGNQV 309
Db 894 GVNLEQVTLDRLLQKVEHSEDNTEKHHQOL---LDLLESVLGNNDNLDSIKTPHTELQ 950
QY 310 QVE-----FATSTANNT-----ANPTPAVDEIKVAKIVLSGLRFGQNTI----- 350
Db 951 KITDHLVKGTTSLANHTNELLGDSLSCLNETTIEDLSLVKLETTGDTSPSKRELPAFSP 1010
QY 351 -----ELSVPTGEGNKNKVPAMTIGNIYLSNENNADKIPCYRRP 389
Db 1011 WTRDSSLIKETTINLSDSKKFVREIYTTSSNQINPEVDV---YDKP 1052

RESULT 10
US-09-268-347-24
; Sequence 24, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
```

US-09-268-347-30

TELECOMMUNICATION INFORMATION:
 REFERENCE/DOC#E1 NUMBER: 1438.01400001/RW
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540

Search completed: June 12, 2002, 10:49:08
Job time: 201 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:47:42 ; Search time 17.44 Seconds
(without alignments)
872.522 Million cell updates/sec

Title: US-09-147-052-2_COPY_64_456

Perfect score: 1998

Sequence: 1 CMSITKKNPNNGQTLEA.....SSNENNADKIPGYRRPCTFL 393

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	154.5	7.7	682	1	NISP_LACLA
2	143	7.2	857	1	AR56_CANAL
3	138.5	6.9	1314	1	SW11_YEAST
4	137	6.9	719	1	YW41_YEAST
5	135.5	6.8	1085	1	CUT7_SCHPO
6	132	6.6	1271	1	Y338_MYCGE
7	131.5	6.6	1164	1	BAG_STRAG
8	129	6.5	1251	1	RBP2_PLAYB
9	128.5	6.4	2334	1	WAPA_BACSU
10	128.5	6.4	3712	1	LMA_DROME
11	128	6.4	995	1	Y109_YEAST
12	127	6.4	956	1	YEF3_YEAST
13	126.5	6.3	444	1	PST1_YEAST
14	126.5	6.3	1630	1	MSPI_PLAFK
15	126.5	6.3	1639	1	MSPI_PLAFW
16	125.5	6.3	1858	1	P3K2_DICDI
17	124.5	6.2	821	1	LIN5_CAEEL
18	124.5	6.2	1010	1	YK1_CAEEL
19	124.5	6.2	1018	1	FNBA_STAAR
20	124	6.2	749	1	MAD1_YEAST
21	124	6.2	750	1	YKS7_YEAST
22	124	6.2	823	1	NSP1_YEAST
23	124	6.2	866	1	WSP1_SCHJA
24	124	6.2	1159	1	N124_SCHPO
25	124	6.2	1433	1	CAT8_YEAST
26	123.5	6.2	537	1	ARP_PLAFA
27	123.5	6.2	1902	1	P2P_LACLC
28	123	6.2	1260	1	ALS1_CANAL
29	123	6.2	1637	1	MSP1_STAAR
30	123	6.2	1790	1	USO1_YEAST
31	122.5	6.1	507	1	FLIC_SALON
32	122	6.1	1117	1	YK96_YEAST
33	122	6.1	1672	1	PMPB_CHLNU

34 121.5 6.1 425 1 YBYO_YEAST
35 121.5 6.1 1140 1 YM96_YEAST
36 121 6.1 491 1 TOLC_SALEN
37 121 6.1 796 1 YS8A_CAEEL
38 121 6.1 901 1 PIF_LACLA
39 121 6.1 979 1 P115_MYCHR
40 121 6.1 1111 1 KIP1_YEAST
41 121 6.1 1902 1 P2P_LACPA
42 120.5 6.0 776 1 VP4_ROTCH
43 120.5 6.0 800 1 INLA_LISMO
44 120.5 6.0 1565 1 PAC_STRMU
45 120 6.0 650 1 PDR4_YEAST

ALIGNMENTS

RESULT 1
NISP_LACLA STANDARD; PRT; 682 AA.
AC Q07596;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nisin leader peptide processing serine protease NISP precursor
DE (EC 3.4.21.-).
GN NISP.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NIZO RS;
RX MEDLINE-93239683; PubMed-8478324;
RA van der Meer J.R., Polman J., Beerthuyzen M.M., Siezen R.J.,
RA Kuipers O.P., de Vos W.M.;
RT "Characterization of the Lactococcus lactis nisin A operon genes
nisp, encoding a subtilisin-like serine protease involved in
precursor processing, and nispR, encoding a regulatory protein
involved in nisin biosynthesis.";
RT J. Bacteriol. 175:2578-2588(1993).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN-6F3;
RX MEDLINE-94213458; PubMed-8161176;
RA Engelke G., Gutowski-Eckel Z., Kiesau P., Siegers K.,
RA Hammelmann M., Entian K.-D.;
RT "Regulation of nisin biosynthesis and immunity in Lactococcus lactis
6F3";
RL Appl. Environ. Microbiol. 60:814-825(1994).
RN [3]
RP 3D-STRUCTURE MODELING.
RX MEDLINE-95357326; PubMed-7630881;
RA Siezen R.J., Rollemans H.S., Kuipers O.P., de Vos W.M.;
RT "Homology modelling of the Lactococcus lactis leader peptidase Nisp
and its interaction with the precursor of the lantibiotic nisin.";
RL Protein Eng. 8:117-125(1995).
CC -!- FUNCTION: CLEAVES THE LANTIBIOTIC NISIN PRECURSOR PEPTIDE.
CC -!- PATHWAY: LAST STEP OF NISIN BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL
(POTENTIAL).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC -!- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS
CC IN THE REGION OF THE MEMBRANE ANCHOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L11061; AAA25200.1; -.
DR EMBL; X76884; CAA54210.1; -.
DR HSP; P29600; IGCI.
DR MEROPS; S08.059; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; FALSE_NEG.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; 1.
KW Hydrolase; Serine protease; Cell wall; Signal; Transmembrane; Zymogen.
FT SIGNAL 1 22
FT PROPEP 23 195
FT CHAIN 196 682
FT FT
FT FT
FT DOMAIN 196 651
FT TRANSMEM 652 679
FT DOMAIN 680 682
FT ACT_SITE 259 259
FT ACT_SITE 306 306
FT ACT_SITE 512 512
FT DOMAIN 652 657
FT CONFLICT 500 500
FT SEQUENCE 682 AA; 74767 MW; D5F29313F2983EC9 CRC64;
Query Match 7.7%; Score 154.5; DB 1; Length 682;
Best Local Similarity 21.0%; Pred. No. 0.11; Mismatches 130; Indels 91; Gaps 13;
Matches 75; Conservative 61;
QY 57 ETYNNLNATLEQKMAKTNLESAINQANTDKTTFONEHPNLVEAYKALKTTLEQRATNL 116
DB 38 ELINHNSNAILSTEGSTTDSINLGAQSPAVKSTTRTE---LDVTGAARKTLLOTSQVOK 93
QY 117 EGLSSTAYNQIRNLVDLTKASSLTKPLDPLNGTLLDSNEITANKNNITLSTINE 176
DB 94 EMKVSQETQVSSE---FSKRDSVTNKEAPVSKDELLEQSEVVVSTSIQKN-KILDN 148
QY 177 OKTNADALSNSFTKKVQIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTV- 235
DB 149 KKRANFVTSPLIKEKPSNKSADGVINDS-----ASPLSYKAKEVVS 193
QY 236 -----WNGDEPSSRILANTNSITDVSITYSIAGTNTKYQFSNYPGS 278
DB 194 LRQPLKNQKVEAQPILLISNSSEKSKASYTNSHDFWDYQW-----DMKY---VTNNGES 243
QY 279 TGYLYFPYKLVKADANNVGLQKLNQVQVEFATISAN-----NT 322
DB 244 YA-LYQPSKKI-----SVGI---IDSGIMEHPDLSLGNFYKNLVPKGGFDNEEPE 293
QY 323 TANTPTAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKNAVPMTG-NIYLLSSNEN 378
DB 294 TGNPSDVID-----KMGHGTAVAGQITANGNLGVAPGITVNIYRFGEN 338
RESULT 2
AR56_CANAL STANDARD; PRT; 857 AA.
AC AR56_CANAL
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE ARG5.6 protein, mitochondrial precursor [Contains: N-acetyl-gamma-
DE glutamyl-phosphate reductase (EC 1.2.1.38) (N-acetyl-L-glutamate
DE semialdehyde dehydrogenase) (NAGSA dehydrogenase); Acetylglutamate
DE kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-glutamate 5-
DE phosphotransferase)].
GN ARG5.6.
```

```
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 64385 / 1001;
RX MEDLINE=97195775; PubMed=9043106;
RA Negredo A., Monteoliva L., Gil C., Pla J., Nombela C.;
RT "Cloning, analysis and one-step disruption of the ARG5,6 gene of
RT Candida albicans.";
RL Microbiology 143:297-302(1997).
CC -!- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
CC + phosphate = N-acetyl-5-glutamyl phosphate + NADPH.
CC -!- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate = ADP + N-acetyl-L-
CC glutamate 5-phosphate.
CC -!- PATHWAY: SECOND AND THIRD STEPS IN ARGinine BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC ACETYLGLUTAMATE KINASE FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAGSA
CC DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X98880; CAA67383.1; -.
DR InterPro; IPR000706; AGPR_act_site.
DR InterPro; IPR001048; Aakinas.
DR InterPro; IPR000534; Semialdh_dh.
DR Pfam; PF00696; aakinas; 1.
DR Pfam; PF01118; Semialdehyde_dhc; 1.
DR Pfam; PF02774; Semialdehyde_dhc; 1.
DR ProDom; PD003765; AGPR_act_site; 1.
DR PROSITE; PS01224; ARG; 1.
DR Oxidoreductase; Transferase; Kinase; Arginine biosynthesis; NADP;
KW Mitochondrion; Multifunctional enzyme; Transit peptide.
FT TRANSIT 1 MITOCHONDRION (POTENTIAL).
FT CHAIN ? ? ACETYLGLUTAMATE KINASE.
FT CHAIN ? 857 N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE
FT REDUCTASE.
FT ACT_SITE 669 669 BY SIMILARITY.
FT SEQUENCE 857 AA; 95066 MW; AE2CEAD8FF8C471 CRC64;
Query Match 7.2%; Score 143; DB 1; Length 857;
Best Local Similarity 23.3%; Pred. No. 0.63;
Matches 99; Conservative 58; Mismatches 134; Indels 134; Gaps 25;
QY 45 IEASLSAY-----SEATVNN---NLNATL-----EOLKMAKTNLESAINQANTD 87
DB 208 VERAINGYLPILTSLAETSSGQLLNVDVAAGELAREFEPLKIVYLNKGGIINGTG 267
QY 88 K-----TTFDEHEPNLV-EAYKALKTTLEQR-----ATNLEGLSTAYNQIRNLVDLYNK 137
DB 268 EKVSAINLDEEYEDLLKESWVYGTUKLKEIHDLQHLPRSSVA-----IIDVNDL 320
QY 138 ASSLITKTLDPLNGTLLDSNEITANKNNIN-NFLSTINEQKTNADALSNSFIKKIQNN 196
DB 321 QKELFTDS-----GAGTL-----IRRGYLLNRNLSRDFG---NPDLRLNALLR----- 361
QY 197 EQSFVGFTTANVQPSNYSFVAFSADVTPVNYKYARTVNGDPPSRILANT----- 249
DB 362 -----DPEIKTKGVSVASVYLKFLDSVQFKS-----YGDPEP-LEVLAIYVEQNDKI 405
QY 250 -----NSITD-----VSMIYSLAGTNTKYQFSNYPGS-PSTGYL 282
DB 406 PKLDEFLSSKTGTLNVTNIDNIFNAIKKDYQSLQVWVNVENDANLFWYFSKDGSAKNGQI 465
```


QY 39 -LQDYAKIEA-----SLSSAYSEATVNNLNA-----TLEQLKMAKTNLESAL 81
 Db 97 LRTSAISAPTQTSOPTTEKLGKLSNRSVNSASNSTTACTGRMLSQLTN-DSPS 155
 QY 82 NOANTDK-----TTFDNEHNLV---EAYKALKTTLPQRATNLEGLSTAY----- 124
 Db 156 NEISTDLQKIFQMDQMSARMEMESFNKLSKIAEQTWVNLKQDNKYKVNKLNIILL 215
 QY 125 -----NOIRNL-VOLYNKASSLTKTLTDLPL--NGSTLDSNEITANKN--- 166
 Db 216 KLVQAPSARPSTNAQKLAIELLSISAVSAYLQKMQNNGSGRQHTADLCTGDSNTHS 275
 QY 167 -----INVTSLTINEQKNA-----DALSNFPIKKVI 193
 Db 276 GINQHRTTNGTIDVNTNTAQNLQNFNSALNTLPDQOHRNNVSNQINSLPNRLGPVI 335
 QY 194 -----QNNQESFV-GTFTANQVPSNYSFVAFSAD----VTPVNYKYARTVWNGDEPS- 242
 Db 336 NTAQANQSQVLHNTHTHOQVNRSPISPNASTDPFKLPNGIKRRRTQSNNAST 395
 QY 243 -----SRILANTNSITDVSWIYS-----LAGTNKYQF-SFSNYGPSTGY 281
 Db 396 NDHASAQAQPIASLPLTNSHSTTSMNTNTHSSIHSGVTSASNFHDLNLSNFGTTTAL 455
 QY 282 LYFPYKLVKAADANNVGLQYKLVNNGVQVEPATSTSANNTTANPTPAVDEIKVAKIVLS 341
 Db 456 SLPSLALDNASFPNQNVPPIINNTQOPLSFSLINQDSTIS----- 498
 QY 342 GLRFGONTIELSYPTGEGNNKVPAMIGNIYSSNENNADKIPGYRRPGT 391
 Db 499 -----EL-LPSCKSGVNT-----NIV-----NNRSTLPSYKPKMT 529

RESULT 5
 CUT7_SCHPO STANDARD; PRT; 1085 AA.
 AC P24339;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 13-JUL-1998 (Rel. 36, Last annotation update)
 DE Kinesin-like protein cut7.
 GN CUT7 OR SPAC25G10.07C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 [1]
 SEQUENCE FROM N.A.
 RP MEDLINE=91015362; PubMed=2145514;
 RX Hagan I., Yanagida M.;
 RA "Novel potential mitotic motor protein encoded by the fission yeast
 RT cut7+ gene.";
 RL Nature 347:563-566(1990).
 RN [2]
 SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA McLean J., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: COULD BE A SPINDLE POLE BODY MOTOR. ON TRANSITION FROM
 CC G2 TO M PHASE OF THE CELL CYCLE, THE SPINDLE POLE BODY DUPLICATES;
 CC THE DAUGHTER POLE BODIES SEED MICROTUBULES WHICH INTERDIGITATE TO
 CC FORM A SHORT SPINDLE THAT ELONGATES TOSPAN THE NUCLEUS AT
 CC METAPHASE. MUTATIONS AT CUT7 BLOCK SPINDLE FORMATION.
 CC -I- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIMC
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch).
 CC EMBL; X57513; CAA40738.1; -.
 DR EMBL; 270691; CAA94636.1; -.
 DR PIR; S14032; S14032.
 DR HSP; P17119; 3KAR.
 DR InterPro; IPR001752; kinesin.
 DR Pfam; PF00225; kinesin; 1.
 DR PRINTS; PRO0380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
 DR PROSITE; PS0067; KINESIN_MOTOR_DOMAIN2; 1.
 KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
 KW Mitosis; Cell cycle; Phosphorylation; Repeat.
 FT DOMAIN 70 435 KINESIN-MOTOR (BY SIMILARITY).
 FT DOMAIN 436 604 COILED COIL (POTENTIAL).
 FT DOMAIN 715 740 COILED COIL (POTENTIAL).
 FT DOMAIN 897 955 COILED COIL (POTENTIAL).
 FT NP_BIND 159 166 ATP (BY SIMILARITY).
 FT REPEAT 987 998
 FT REPEAT 999 1010
 FT MOD_RES 1011 1011
 FT CONFLICT 34 61 PHOSPHORYLATION (BY CDC2) (BY
 FT SIMILARITY).
 FT SASNPKRREPPTIDTGYPDSDTNSPT ->
 FT LRAILGNVSLLLTL (IN REF. 1).
 FT SEQUENCE 1085 AA; 122133 MW; 5669277875559D58 CRC64;
 Query Match 6.8%; Score 135.5; DB 1; Length 1085;
 Best Local Similarity 20.9%; Pred. No. 2.2;
 Matches 97; Conservative 76; Mismatches 182; Indels 109; Gaps 21;

QY 13 NGQ-TOLEARMELTDLINAKA-WTLASL----QDYAKIEASLSSAYSEATVNNLNA 66
 Db 623 NGYFTLLDNDFASMEELLNTHSNQLLISMTKITEHQSLDEALQSARSSCAVPNSLDLI 682
 QY 67 LEOLKMAKTNLESAINQANTDKTTFDNEHNP-----LVEAYKALKTT---LEQATNLEG 118
 Db 683 VSELKDSKNSLLDALHSLQDLSMSQKLGNGISSLELQDKMKESYQLVQELSLYN 742
 QY 119 LSTAYNQIRNNLVLYNKASSLITKTLPLNGTLLDSEIT-----TA 163
 Db 743 LQHTHEESQKELMYGVNRDIDALVKTCTSLNADIILSDYISDQSKFESKODLIANI 802
 QY 164 NKINNTLSTINEQ-KTNADAL-----SNSFIKK---VIONNEQSFVGTFTNANVQPSN 213
 Db 803 GKIVSNFLOEQNESLYTKADILHSHLNDNSIRKANEIMNRSSEFL----RNA----- 853
 QY 214 YSFVAFSADVTPVNYKYARTVWNG-----DEPSRLANTNSITDVSWIYSLA----- 262
 Db 854 ----ASQAEIVGANKERIKQTVENGSQLDLSKSKATHSNRSMYD----HCLALAESQK 905
 QY 263 GINTKYQ-----SFSNYGPGSTGYLYFPYKLVKAADANNVGL--QYKLVNNGVQ 309
 Db 906 GVNLEVQTLDRLLQKYKHESEDNTEKHHQOL---LDLLESVGNNDNLDSIKTPHTELQ 962
 QY 310 QVE---FATSTSANNTT-----ANPTPAVDEIKVAKIVLSGLRFGONTI----- 350
 Db 963 KIYDHLVKGTTSLANTHTNELLGLGDSLCNLETTIEDTSLVKLETTGDTPSKRELPA 1022
 QY 351 -----ELSVPTGEGNNKVPAMIGNIYSSNENNADKIPGYRRP 389
 Db 1023 WTRDSSLIKETTLLNLDSDKKFVREYTSSTNQTNPDV--YDKP 1064
 RESULT 6
 Y338_MYCGE
 ID Y338_MYCGE STANDARD; PRT; 1271 AA.
 AC P47580;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical lipoprotein MG338 precursor.

Query Match 6.6%; Score 131.5; DB 1; Length 1164;
Best Local Similarity 22.8%; Pred. No. 3.9;
Matches 84; Conservative 51; Mismatches 157; Indels 77; Gaps 15;

QY 6 KIDAMPNGQTOLEA-ARWELTDLINA-KAMTLASLDYAKIEASLSAYSE- - - - -AET 58
DB 230 RKQAOADKDEAEVREELGKLFSSKAGLDQBIQEHVKKETSSEENTQKVDEHYANS 289
QY 59 VNNLNATLEQLKMAKTNLESAINOANTDKTTFDNEHPNL- - - - -VEAYKALKT 107
DB 290 LONLAQKSLLEEDKATN- - - - -EATQVKNQFLENAQKLEIQPLIKETNVKLYKANSE 344
QY 108 TLEQATNLEGLSSSTAYNOIRNLDVLYNKASLITKTLDPLNGGTLTDSNEITANKNI 167
DB 345 SLEQVEKELKHNSA- - - - -NLEDVAKSKEIVREYEGKLNQSKNLP- - - - -ELKOLEEBA 395
QY 168 NWTILSTINEQ- - - - -KTNDALNSFIKVI- - - - -ONNEQSFVGTFTNANVQSNYSFVAF 219
DB 396 HSKLQOVDFRKKFTSEQVTPKKVRKRDLAANNQOKI- - - - -ELTVSPENI- - - - - 445
QY 220 SADVTVPVYKARRTVWNGDEPSSRLANTNSITDVSIVSLAGTNTKYQFSFS- - - - -N 274
DB 446 - - - - -TYEGEDVKFTAKSDSKTTLDFSDLL- - - - -TKNPSVSDRISN 487
QY 275 YGSGTG- - - - -YLYFPYKLVKAADANNVGLQYKLNNGVQVFEATSTANP-TPAV 330
DB 488 YKNTDNHIAETIKNLKLNESQTVTLKAKDDSGNVVEKTEITITVQKKEKQVPTPEQ 547
QY 331 DEIKVAKIV 339
DB 548 KDSKTEEV 556

RESULT 8
RBP2_PLAVB STANDARD; PRT; 1251 AA.

ID RBP2_PLAVB AC Q00799;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Reticulocyte binding protein 2 (Fragment).
GN RBP2.
OS Plasmodium vivax (strain Bellem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92315338; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
merozoites".
RL Cell 69:1213-1226(1992).
CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC HUMAN RETICULOCYTE CELLS.
CC -1- SUBCELLULAR LOCATION: Membrane-bound (Probable).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M88098; AAA29744.1; -
KW Malaria; Receptor; Membrane.
FT NON_TER 1
FT NON_TER 1251 1251
SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;

Query Match 6.5%; Score 129; DB 1; Length 1251;
Best Local Similarity 21.1%; Pred. No. 5.9;

Matches 89; Conservative 70; Mismatches 140; Indels 122; Gaps 20;

QY 15 QTOLEAARMEITDLINAKAMTLASLDYAKIEASLSAYSEAB- - - - -TVNNLNAT 66
DB 424 ETNLESVHNLSIEITNIK- - - - -QGGEKI- - - - -YSKAKIDIMQIKATSENTAETK 469
QY 67 LEOLKMAKTNLESATNOANTDKTTFDNEHPNLVEAYKALKTTLEQATNLEGLSSSTAYNQ 126
DB 470 LEKVKKDDQSNVYVNLQITTER- - - - -NLIVTEKRNLCIDSTIINIEG- - - - -ALKE 517
QY 127 IRNNL- - - - -VDLYNKASSLITKTLDPLNGGTLTDSNEIT- - - - -TANKN 166
DB 518 SKGNEYGFLEKLEEGKRNKLVKITKKS- - - - -INSTVG- - - - -NFSLENNFDLNOYDFENK 572
QY 167 INNTLSTINEQTNADALNSP- - - - -IKKVIQNNQSFVGTFTNANVQSNYSFVAFSADV 223
DB 573 INDYENKMG- - - - -IYNEFEGLINKISEN- - - - -LRNASENTSIDYN- - - - -SAKT 613
QY 224 TPNVYKARRTVWNGDEPSSRLANTNSITDVSIV- - - - -SLAGTN- - - - -TKYQFSFSNYG 276
DB 614 LRLEAQKEKVNLLNKEEENKYLROVKVYSEFRFFENKESLDKINEMIKKQLTIVNEGH 673
QY 277 PSTGVLFPYKLVKAADANNVGLQYKLNNGVQVFEATSTANN- - - - -TTA- - - 324
DB 674 GNVKQLVENIK- - - - -ELVDENNLSDLKQATGKNEEIQKITHSTLKNKAKTILGHVDTSAKY 731
QY 325 - - - - -NPTPAVDE- - - - -IKVAK- - - - -IYLSGLRFGONTIELSVPTCEGNNKV 364
DB 732 VGIKIPPELALTELLGDAKLTAQELKFESKNNVLETTNNMKNKTNLDVHKNQIDAYKV 791
QY 365 A 365
DB 792 A 792

RESULT 9
WAPA_BACSU STANDARD; PRT; 2334 AA.

ID WAPA_BACSU AC Q07833;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Wall-associated protein precursor.
GN WAPA OR N17G.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=93302506; PubMed=8316082;
RA Foster S.J.;
RT "Molecular analysis of three major wall-associated proteins of a
RT Bacillus subtilis 168: evidence for processing of the product of a
RT gene encoding a 258 kDa precursor two-domain ligand-binding
RT protein".
RL Mol. Microbiol. 8:299-310(1993).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=95219088; PubMed=7704263;
RA Yoshida K.-I., Sano H., Seki S., Oda M., Fujimura M., Fujita Y.;
RT "Cloning and sequencing of a 29 kb region of the Bacillus subtilis
RT genome containing the hut and wapa loci".
RL Microbiology 141:337-343(1995).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=97124196; PubMed=8969509;
RA Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA Miwa Y., Fujita Y.;
RT "Sequencing of a 65 kb region of the Bacillus subtilis genome

containing the lic and cel loci, and creation of a 177 kb contig
RT covering the gnt-sacxy region.";
RL Microbiology 142:3113-3123(1996).
CC -|- FUNCTION: STILL UNKNOWN. NOT INVOLVED IN CELL MEMBRANE METABOLISM,
CC MOTILITY, SECRETION OR DIFFERENTIATION.
CC -|- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MAY BE RELEASED
CC INTO THE MEDIUM.
CC -|- DOMAIN: HAS TWO LIGAND-BINDING DOMAINS; THE N-TERMINUS, HAS THREE
CC 101 AA REPEATS WHICH ARE RESPONSIBLE FOR CELL WALL BINDING; THE
CC C-TERMINUS CONSISTS OF TWO BLOCKS OF RESIDUES WITH A CONSERVED
CC MOTIF REPEATED 31 TIMES.
CC -|- SIMILARITY: THE REPEATED MOTIF IN THE C-TERMINUS HAS SOME
CC SIMILARITY TO THE REPEAT IN E. COLI RHS GROUP OF PROTEINS (RHS-A-D).
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L05634; AAA22883.1; -;
DR EMBL; D31856; BAA06656.1; -;
DR EMBL; D29985; BAA06260.1; -;
DR EMBL; D83026; BAA11683.1; -;
DR EMBL; Z99124; CAB15959.1; -;
DR PIR; S32920; S32920.
DR Subtilist; BGI0797; wapa.
DR InterPro; IPR003305; CBD_6.
DR Pfam; PF02018; CBD_6; 1.
DR Cell wall; Repeat; Signal; Complete proteome.
KW SIGNAL 1 28
FT CHAIN 29 2334
FT OR 32 (POTENTIAL).
FT WALL-ASSOCIATED PROTEIN.
FT DOMAIN 504 869
FT REPEAT 504 605
FT REPEAT 636 736
FT REPEAT 769 869
FT DOMAIN 1021 2139
FT
FT 31 X 21 AA APPROXIMATE TANDEM REPEATS OF
FT X(4)-G-X(4)-[YF]-X-D-X(2)-G-X(4).
FT 2-1.
FT 2-1.
FT 2-2.
FT 2-3.
FT 2-4.
FT 2-5.
FT 2-6.
FT 2-7.
FT 2-8.
FT 2-9.
FT 2-10.
FT 2-11.
FT 2-12.
FT 2-13.
FT 2-14.
FT 2-15.
FT 2-16.
FT 2-17.
FT 2-18.
FT 2-19.
FT 2-20.
FT 2-21.
FT 2-22.
FT 2-23.
FT 2-24 (APPROXIMATE).
FT 2-25.
FT 2-26.
FT 2-27.
FT 2-28.
FT 2-29.
FT 2-30.
FT 2-31.
FT 2334 AA; 258329 MW; B75138CCD278BAA3 CRC64;
SQ SEQUENCE

Query Match 6.4%; Score 128.5; DB 1; Length 2334;
Best Local Similarity 22.18; Pred. No. 14;
Matches 91; Conservative 56; Mismatches 152; Indels 113; Gaps 22;
QY 3 SITKDDANPNNGOTOLEAARM-ELTDLINAKAMTILASLDYAKIEASLSAYSEATVN- 60
DB 1772 SWANKEQNTTKRTPDNKNRNLTELTDRGSGQTWTPSPSDKLK---TFSWIHGDKGTNQ 1828
QY 61 ---NNLNATLSQLKMAKT---NLESAINQANTDKTTFDNEHNLV----- 99
DB 1829 FTYNKLDQMIEMKOSTSSYSFDYDENGNVQTFIT-GNGGGTSFSDERNLVSSLHIGDKN 1887
QY 100 -----EAYK----ALKTTLEQRATNLEGLSSTAYNQLRNLDVLYNKASSLITTTLDPL 149
DB 1888 GGDILTESYEDVANDGNRTTISSAS---GKVOYEGKLNQLVKETHEDGVITYYDGF 1943
QY 150 NG---GYLLD-SNEITTPANKNNITLSTINEQKTNADALSNSFIKKVIONNEQSFVGT 204
DB 1944 GNRKTVTTIKDGSSKTVAASFNIMNQLTKVANDESISYD-----KNGNRSTDGKF 1992
QY 205 TNANVQPSNYSFVAFSADVTP-VNKKYARRIVWNGDEPSSRILANTNSITDVSWIYSLAG 263
DB 1993 TYTWDADNLTAVTKKGDCKPFATYKY-----DEKGNRIQKTVNGKV-TNYFYDGDS 2043
QY 264 TNTKYQFSFN-----YGPS-----TGILYFPPYKLVKAADANNVGLQYKLN-NGN 307
DB 2044 LNVLYETDADNNVTKSYTYGSGQLLSTYENGKKYF-----YHNAHGD 2087
QY 308 VOOVEFATSTSANNTTA-----NP--TPAVDEIKVAKIVLSGLRFGONT 349
DB 2088 I-----IAISDSTGKTVAKYQYDAMGNPTKTEASDEVKONRYRAGYQYDEET 2135
RESULT 10
LMA_DROME
ID LMA_DROME STANDARD; PRT: 3712 AA.
AC Q00174;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Laminin alpha chain precursor.
GN LANA OR LAMA.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93049203; PubMed=1425586;
RA Kusche-Gullberg M., Garrison K., Mackrell A.J., Fessler L.I.,
RA Fessler J.H.;
RT "Laminin A chain: expression during Drosophila development and
RT genomic sequence.";
RL EMBO J. 11:4519-4527(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=94038678; PubMed=8223265;
RA Henchcliffe C., Garcia-Alonso L., Tang J., Goodman C.S.;
RT "Genetic analysis of laminin A reveals diverse functions during
RT morphogenesis in Drosophila";
RL Development 118:325-337(1993).
RN [3]
RP SEQUENCE OF 1762-3712 FROM N.A.
RX MEDLINE=92078147; PubMed=1744083;
RA Garrison K., Mackrell A.J., Fessler J.H.;
RT "Drosophila laminin A chain sequence, interspecies comparison, and
RT domain structure of a major carboxyl portion.";
RL J. Biol. Chem. 266:22899-22904(1991).
CC -|- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF

CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
 CC -|- FUNCTION: DIVERSE FUNCTIONS DURING MORPHOGENESIS IN DROSOPHILA.
 CC COMPLETE LOSS-OF-FUNCTION MUTATIONS LEAD TO LATE EMBRYONIC
 CC LETHALITY. CERTAIN PARTIAL LOSS-OF-FUNCTION MUTATIONS GIVE RAISE
 CC TO ESCAPER ADULTS, WHICH HAVE ROUGH EYES ASSOCIATED WITH CHANGES
 CC IN CELL FATE AND PATTERN, MISSHAPEN LEGS AND DEFECTS IN WING
 CC STRUCTURE.
 CC -|- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
 CC -|- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
 CC MEMBRANES (MAJOR COMPONENT).
 CC -|- TISSUE SPECIFICITY: NEWLY FORMED MESODERM AND LATER PROMINENTLY
 CC EXPRESSED IN HEMOCYTES, WHICH ALSO SYNTHESIZE COLLAGEN IV.
 CC -|- DEVELOPMENTAL STAGE: DURING MORPHOGENESIS, MOSTLY IN EMBRYO
 CC DEVELOPMENT AT 10-12 HOURS.
 CC -|- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -|- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
 CC -|- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -|- SIMILARITY: CONTAINS 21.5 LAMININ EGF-LIKE DOMAINS.
 CC -|- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV (DOMAIN IV) IS NOT
 CC SIMILAR TO LAMININ DOMAIN IV.
 CC -|- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; M96388; AAA28662.1; -;
 CC DR EMBL; L07288; AAC37178.1; -;
 CC DR EMBL; W75882; AAA28661.1; -;
 CC DR HSSP; P02468; ITLE.
 CC DR FlyBase; FBgn0002526; LanA.
 CC DR InterPro; IPR000561; EGF-like.
 CC DR InterPro; IPR001886; LamNT.
 CC DR InterPro; IPR000034; Laminin_B.
 CC DR InterPro; IPR002049; Laminin_EGF.
 CC DR InterPro; IPR001791; Laminin_G.
 CC DR InterPro; IPR001230; Prenyltn.
 CC DR Pfam; PF00052; laminin_B; 1.
 CC DR Pfam; PF00053; laminin_EGF; 20.
 CC DR Pfam; PF00054; laminin_G; 5.
 CC DR Pfam; PF00055; laminin_Nterm; 1.
 CC DR PRINTS; PRO00011; EGFLAMININ.
 CC DR ProDom; PD002082; LamNT; 1.
 CC DR ProDom; PD003031; Laminin_B; 1.
 CC DR SMART; SM00180; EGF_Lam; 17.
 CC DR SMART; SM00001; EGF_Like; 1.
 CC DR SMART; SM00281; LamB; 1.
 CC DR SMART; SM00282; LamG; 5.
 CC DR SMART; SM00136; LamNT; 1.
 CC DR PROSITE; PS00022; EGF_1; 17.
 CC DR PROSITE; PS01186; EGF_2; 5.
 CC DR PROSITE; PS01248; LAMININ_TYPE_EGF; 19.
 CC DR PROSITE; PS50025; LAM_G_DOMAIN; 5.
 CC KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
 CC Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 CC FT SIGNAL 1 22 POTENTIAL.
 CC FT CHAIN 23 3712 LAMININ ALPHA CHAIN.
 CC FT DOMAIN 25 272 LAMININ N-TERMINAL (DOMAIN VI).
 CC FT DOMAIN 273 332 LAMININ EGF-LIKE 1.
 CC FT DOMAIN 333 402 LAMININ EGF-LIKE 2.
 CC FT DOMAIN 403 447 LAMININ EGF-LIKE 3.
 CC FT DOMAIN 448 494 LAMININ EGF-LIKE 4.
 CC FT DOMAIN 495 540 LAMININ EGF-LIKE 5.
 CC FT DOMAIN 541 586 LAMININ EGF-LIKE 6.

FT DOMAIN 587 631 LAMININ EGF-LIKE 7.
 FT DOMAIN 632 676 LAMININ EGF-LIKE 8.
 FT DOMAIN 677 731 LAMININ EGF-LIKE 9.
 FT DOMAIN 732 784 LAMININ EGF-LIKE 10.
 FT DOMAIN 785 815 LAMININ EGF-LIKE 11 (INCOMPLETE).
 FT DOMAIN 816 1374 DOMAIN IV'.
 FT DOMAIN 1375 1420 LAMININ EGF-LIKE 12.
 FT DOMAIN 1421 1465 LAMININ EGF-LIKE 13.
 FT DOMAIN 1466 1513 LAMININ EGF-LIKE 14.
 FT DOMAIN 1514 1564 LAMININ EGF-LIKE 15.
 FT DOMAIN 1565 1775 LAMININ EGF-LIKE 16 (N-TERMINAL).
 FT DOMAIN 1776 1808 LAMININ EGF-LIKE 17 (C-TERMINAL).
 FT DOMAIN 1809 1858 LAMININ EGF-LIKE 18.
 FT DOMAIN 1859 1916 LAMININ EGF-LIKE 19.
 FT DOMAIN 1917 1969 LAMININ EGF-LIKE 20.
 FT DOMAIN 1970 2016 LAMININ EGF-LIKE 21.
 FT DOMAIN 2017 2063 LAMININ EGF-LIKE 22.
 FT DOMAIN 2064 2111 DOMAIN II AND I.
 FT DOMAIN 2112 2671 LAMININ G-LIKE 1.
 FT DOMAIN 2672 2868 LAMININ G-LIKE 2.
 FT DOMAIN 2869 3048 LAMININ G-LIKE 3.
 FT DOMAIN 3055 3223 LAMININ G-LIKE 4.
 FT DOMAIN 3249 3528 LAMININ G-LIKE 5.
 FT DOMAIN 3534 3709 COILED COIL (POTENTIAL).
 FT DOMAIN 3710 2249 COILED COIL (POTENTIAL).
 FT DOMAIN 2301 2321 COILED COIL (POTENTIAL).
 FT DOMAIN 2376 2450 COILED COIL (POTENTIAL).
 FT DOMAIN 2541 2676 COILED COIL (POTENTIAL).
 FT DOMAIN 3270 3296 POLY-THR.
 FT DISULFID 273 282 BY SIMILARITY.
 FT DISULFID 275 296 BY SIMILARITY.
 FT DISULFID 298 307 BY SIMILARITY.
 FT DISULFID 310 330 BY SIMILARITY.
 FT DISULFID 333 342 BY SIMILARITY.
 FT DISULFID 335 367 BY SIMILARITY.
 FT DISULFID 370 379 BY SIMILARITY.
 FT DISULFID 382 400 BY SIMILARITY.
 FT DISULFID 403 414 BY SIMILARITY.
 FT DISULFID 405 421 BY SIMILARITY.
 FT DISULFID 423 432 BY SIMILARITY.
 FT DISULFID 435 445 BY SIMILARITY.
 FT DISULFID 448 460 BY SIMILARITY.
 FT DISULFID 450 468 BY SIMILARITY.
 FT DISULFID 470 479 BY SIMILARITY.
 FT DISULFID 482 492 BY SIMILARITY.
 FT DISULFID 495 507 BY SIMILARITY.
 FT DISULFID 497 514 BY SIMILARITY.
 FT DISULFID 516 525 BY SIMILARITY.
 FT DISULFID 528 538 BY SIMILARITY.
 FT DISULFID 541 553 BY SIMILARITY.
 FT DISULFID 543 560 BY SIMILARITY.
 FT DISULFID 562 571 BY SIMILARITY.
 FT DISULFID 574 584 BY SIMILARITY.
 FT DISULFID 587 599 BY SIMILARITY.
 FT DISULFID 589 605 BY SIMILARITY.
 FT DISULFID 607 616 BY SIMILARITY.
 FT DISULFID 619 629 BY SIMILARITY.
 FT DISULFID 632 644 BY SIMILARITY.
 FT DISULFID 634 650 BY SIMILARITY.
 FT DISULFID 652 661 BY SIMILARITY.
 FT DISULFID 664 674 BY SIMILARITY.
 FT DISULFID 677 691 BY SIMILARITY.
 FT DISULFID 679 700 BY SIMILARITY.
 FT DISULFID 702 711 BY SIMILARITY.
 FT DISULFID 714 729 BY SIMILARITY.
 FT DISULFID 732 746 BY SIMILARITY.
 FT DISULFID 734 753 BY SIMILARITY.
 FT DISULFID 755 764 BY SIMILARITY.
 FT DISULFID 767 782 BY SIMILARITY.
 FT DISULFID 1375 1387 BY SIMILARITY.
 FT DISULFID 1377 1394 BY SIMILARITY.
 FT DISULFID 1396 1405 BY SIMILARITY.


```

FT DISULFID 1408 1418 BY SIMILARITY.
FT DISULFID 1421 1429 BY SIMILARITY.
FT DISULFID 1423 1436 BY SIMILARITY.
FT DISULFID 1438 1447 BY SIMILARITY.
FT DISULFID 1450 1463 BY SIMILARITY.
FT DISULFID 1466 1480 BY SIMILARITY.
FT DISULFID 1468 1487 BY SIMILARITY.
FT DISULFID 1489 1498 BY SIMILARITY.
FT DISULFID 1501 1511 BY SIMILARITY.
FT DISULFID 1514 1526 BY SIMILARITY.
FT DISULFID 1516 1533 BY SIMILARITY.
FT DISULFID 1535 1544 BY SIMILARITY.
FT DISULFID 1547 1562 BY SIMILARITY.
FT DISULFID 1859 1874 BY SIMILARITY.
FT DISULFID 1861 1885 BY SIMILARITY.
FT DISULFID 1887 1896 BY SIMILARITY.
FT DISULFID 1899 1914 BY SIMILARITY.

Query Match 6.4%; Score 128.5; DB 1; Length 3712;
Best Local Similarity 23.4%; Pred. No. 24;
Matches 91; Conservative 48; Mismatches 155; Indels 95; Gaps 19;

Qy 26 TDLINAKAMTILASLDYAKIEASISAYSEATVNNLNATLEQLKAKTNTLESAIQ-- 83
Db 2502 TDLQARQSLQKQD--DLEPLRNASAGVKOKISAVNNATEHQLK----DINKLIDQLP 2555
Qy 84 ANTDKTTFDNEHPNLVEAYKAKTTLLEORATNLEGLSTAYNQIR-----NNLVLDLYNKA 138
Db 2556 AESORDMKNSANASDALEILKNVLE-----ILEPVSQVOTPEKLEKAHGINRDLDTNKD 2611
Qy 139 SSLITKTLPLNGGTLDSNEITTANKNNINTLSTINEQKTNADALNSFIKKVIQNNQ 198
Db 2612 VSQANKQLDDVEGSV-----SKLNELAEIDIEQQHVRGVSQSQRLGOEI----- 2654
Qy 199 SFVGTFTNANQPSNYSFVAFSADYTPVNYKYARTVWNGDEPSSRILANTNSITDYSWI 258
Db 2655 -----ENLKAQVAAQRLANSIKV-GVNFK-----PSTILELKTPEKT----- 2691
Qy 259 YSLAGTNTKYQFSNYPGSPSTGYLYFPYKLVKAADANNVGLQYKLVNNGVQVFEATSTS 318
Db 2692 -KLLATRLNLTSTYFTTEPS-GFLLYLGNDNKTAQKNDFVAVEIVNG-----YPLTIDL 2745
Qy 319 ANNTTANPTPAVDEIKVAKIVLSGL-----RFGON---TIELSVPTGEGNMKNKVPAMI 368
Db 2746 CNGP-----ERITSDKYVADGRWYQAVVDRMGPNAKLTIRELPNGD-----VVEHS 2792
Qy 369 GNYLSSNEN--NADK-----IPGYRRPG 390
Db 2793 KSGYLEGSONILHYDKNSRLRVGQY--PG 2819

RESULT 11
YIQ9_YEAST STANDARD; PRT; 995 AA.
AC P40442;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hypothetical 99.7 kDa protein in Sdi1 5' region precursor.
GN YIL169C OR Y19402.07C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Barrall B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;

```

```

RL Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z46921; CAA87023.1; --
DR SGD; S0001431; YIL169C.
DR InterPro; IPR004089; Chemotaxis_transducer.
DR InterPro; IPR000727; T_SNARE.
KW Hypothetical protein; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 995 HYPOTHETICAL PROTEIN YIL169C.
FT CARBOHYD 28 28 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 35 35 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 664 664 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 995 AA; 99735 MW; F63E287A03F137EC CRC64;

Query Match 6.4%; Score 128; DB 1; Length 995;
Best Local Similarity 16.1%; Pred. No. 5;
Matches 73; Conservative 100; Mismatches 152; Indels 128; Gaps 15;

Qy 3 SITKDDANPNNGQTOLEAARMELTDLINAKAMTILASLDYAKIEASISAYSEATVNNN 62
Db 136 SVSOSSSASDVSSVSQSSSSASDVSSVSQSSASD-----VSSVSQSSASDSD 188
Qy 63 LNAYLEOLKMAKTNTLESAINQANTDKTTFDNEHPNLVEAYKAKTTLLEORATNLEGLST 122
Db 189 VSSVSQSSASDVSSVSQSSSSASDVSSVSQSSASDSDVSSVSQSSASSTSGVSSS 248
Qy 123 AYNOIRNLVLYNKASSLITKTLPLNGGTLDSNEITTANKNNINTLSTINEQKTNAD 182
Db 249 GSQSVSS-----ASGSSSSFPQST-----SSASTASGSAATSNLSSTSSASAS 293
Qy 183 ALSNSFIKK-----VIONEQSF---VG 202
Db 294 ATASNLSSSSDGTIYLPPTTISGDLTLFGKVIATEGVVVAAGAKTLTLDGKYSFADLK 353
Qy 203 TFTNANQPSNYSFVAFSADYTPVNYKYARTVWNGDE---PSSRILANTNSITDYSWI 259
Db 354 VYGLLVKSKRETYPGTEFDISGENFDVTGN--FNAESAATSAISIYFTSSPNSGDI 411
Qy 260 SLA-GTNTKYQFSNYPGSPSTGYLYFPYKLVKAADANNVGLQYKLVNNGVQVFEATST- 317
Db 412 SLSLSKSKKGEVTFSPYNSGAFSF-----SNAI-----LNGGSVGLQRRDDTE 456
Qy 318 -SANTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNMKNKVPAMIGNIYL--- 373
Db 457 GSVNNGEIN-----LDNGSTYIVPEVPVSGKGTVNIIS---GNLYLHYP 496
Qy 374 -----SSNENADKIP--GY 386
Db 497 DFTFGQTVVFKGEGVLAVDPTETNATPIPVVGY 529

RESULT 12
YEF3_YEAST STANDARD; PRT; 956 AA.
AC P32618;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 106.1 kDa protein in GLY1-GDAl intergenic region.
GN YEL043W OR SVGP-ORF14.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

```

OX NCBI_TaxID=4932;
 RP [1]
 RA SEQUENCE FROM N.A.
 RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,
 RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.,
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RA SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman C., Lashkari D., Lew H., Lin D.,
 RA Hyman R., Kayser A., Komp C., Namath A., Norgren R., Oefner P., Oh C.,
 RA Mesedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yelton M., Bolstein D., Davis R.W.,
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 RP [3]
 RA FIBRONECTIN TYPE III DOMAIN.
 RX MEDLINE=97148176; PubMed=8994808;
 RA Bateman A., Chothia C.;
 RT "Fibronectin type III domains in yeast detected by a hidden Markov
 model";
 RL Curr. Biol. 6:1544-1546 (1996).
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U18779; AAB64999.1; -
 DR PIR: S30834; S30834.
 DR SGD: S0000769; YEL043W.
 DR InterPro: IPR003961; FN_III.
 DR Pfam: PF00041; fn3; 1.
 DR SMART: SM00060; FN3; 1.
 KW Hypothetical protein.
 FT DOMAIN 35 125 FIBRONECTIN TYPE-III.
 SQ SEQUENCE 956 AA; 106132 MW; 3F78B09A0FCA03AF CRC64;
 Query Match 6.4%; Score 127; DB 1; Length 956;
 Best Local Similarity 19.7%; Pred. No. 5.4;
 Matches 104; Conservative 86; Mismatches 163; Indels 176; Gaps 25;
 QY 4 ITKKNPNNGQQLAARMLTDL-----INA-KAMTLASLDYAKTEASLSAYSEA 56
 DB 337 LNESASVANINKEIESLQNEISKMEESKRLNASKSLITSVNANVENDKPIASGEL 396
 QY 57 ETYNNLN-ATLE-----OLKMAKTNLESAINQANTDKTFONEHPNLVEAYKALK 106
 DB 397 SAVLKLNDFTEKNGFLSNAGEEFLSKNLSLTKMTKQELSIDQE---LEANKLOR 453
 QY 107 TTLEQATNLEGLSTAYNQIRNNLVLYNKASSLITKTL-DPL--NGTLLDSNEITTA 163
 DB 454 SNLLKKSIALE-----NQF--NENSLNRR--NLTKLMVQPKYKNGDSLAATNSNSA 502
 QY 164 NKN-----TNNTLTINEOKTNADAL----- 184
 DB 503 EKNRSGSGIQLPLSNNSRTSGIDLSNNKSN--NSNADSAPPLRLHPVSPSPNEP 560
 QY 185 ---SNSFIKKVIONNE--QSPVGGTFWAN---VQPSNPFVA-----PSADVTVPNYK 229
 DB 561 IQPSSLSLQTDQTDNRSLNLSHISNNENKQPSYSHALPTTATANATATATN-G 619
 QY 230 YARTTVWNG---DEPSSRIILA-----NTNSITDVSWIYSLAGTNTKYQFSFNYG 276
 DB 620 HSRSLNLTWTAQFAQPSHQYQVSTELDQAFEDNANHL--ISGLQNMIVDETDYDNLISNYS 677

QY 277 -----PSTGYLYPPYKLVKAADANNVGLQYKLANGNQV 309
 DB 678 KGFTTDELNDYWTKOQPOVRSTNESLFTTGTPTMSYK-----ANPVISPT--SSSHLR 729
 QY 310 QVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRF-----GONTIELS-- 353
 DB 730 QTSNATNTPMHPQSLAATLNDPSLQSFVRSYSGYSAPQANSLONNINGNETENISPR 789
 QY 354 -----VPTGEGNMKNVAPMIGNIYLSNENNAKDI 393
 DB 790 ISSDFENLLVPNLSPRLSNDVPIVPGNNNTLTTPSHSNILTMNHQPTADNI 838
 RESULT 13
 PSTL_YEAST
 ID PSTL_YEAST STANDARD: PRT: 444 AA.
 AC Q12355;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Protoplast secreted protein 1 precursor.
 GN PST1 OR YDR055W OR D4214 OR YD9609.09.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96381250; PubMed=8789263;
 RA Brandt P., Ramlow S., Otto B., Bloecker H.;
 RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm
 RT of Saccharomyces cerevisiae chromosome IV";
 RL Yeast 12:85-90 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP IDENTIFICATION.
 RX MEDLINE=99251092; PubMed=10234784;
 RA Pardo M., Monteoliva L., Pla J., Sanchez M., Gil C., Nombela C.;
 RT "Two-dimensional analysis of proteins secreted by Saccharomyces
 RT cerevisiae regenerating protoplasts: a novel approach to study the
 RT cell wall";
 RL Yeast 15:459-472 (1999).
 RN [4]
 RP GPI-ANCHOR.
 RX MEDLINE=20469049; PubMed=11016834;
 RA Terashima H., Yabuki N., Arisawa M., Hamada K., Kitada K.;
 RT "Up-regulation of genes encoding glycosylphosphatidylinositol
 RT (GPI)-attached proteins in response to cell wall damage caused by
 RT disruption of FKS1 in Saccharomyces cerevisiae";
 RL Mol. Gen. Genet. 264:64-74 (2000).
 CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR AND
 CC SECRETED BY REGENERATING PROTOPLASTS.
 CC -!- SIMILARITY: BELONGS TO THE SPS2 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X84162; CAA58971.1; -
 DR EMBL: Z74351; CAA98873.1; -
 DR EMBL: Z49209; CAA89084.1; -
 DR SGD: S0002462; PST1.
 DR COMPLEYEAST-2DPAGE; Q12355;
 KW Glycoprotein; Membrane; GPI-anchor; Signal.
 FT SIGNAL 1 19 POTENTIAL.

FT CHAIN 20 ? PROTOPLAST SECRETED PROTEIN 1.
FT PROPEP ? 444 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 356 416 SER-RICH.
FT CARBOHYD 57 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 444 AA; 45776 MW; 230F60CACA5921A4 CRC64;

Query Match 6.3%; Score 126.5; DB 1; Length 444;
Best Local Similarity 18.1%; Pred. No. 2.3; Indels 105; Gaps 15;
Matches 73; Conservative 79; Mismatches 146; Indels 105; Gaps 15;

Qy 16 TQLEAARME-LTDLINAKAMTASLDQYAKTEASLSAYSEATVNNLNATLQKMAK 74
Db 88 TSFAADLESITSDLSQSLIITL-----SASFGSLQSVDSIKLITLPAISSFT 136
Qy 75 TNLSAINQANTDKTDENEHPNLVEAYKALKTTLEQRATNLEGLSS--TAYNIQIRNLV 132
Db 137 SNIKSANNIYSDTSL-----QSDGFSALKVNVFNVNNKKLTSIKSPVETVSDSLQ 190
Qy 133 DLYNKASSLIPTKLDPL--NGCTLLDSNEITANKN-----INNTLSINEOKTNA 181
Db 191 FSNQNGOTKI--TDDDLWANNISLTDVHSFANLQKINSLSLFINNLSISLFTKLNT 248
Qy 182 ----DALSNSFIKK-----VIQNN-----EQSFVGTFTN 206
Db 249 IGQTFISVNDYLNKLSFNLSLTGGALVANNGLQKIGGLDNLTTIGGTLEVVGNFTS 308
Qy 207 ANVQPSNYSFVAFADVTVPVNYKARTVWNGDEPSSRLANTNSITDVSIIYSLAGTNT 266
Db 309 LNL--DSLKSVKGGADVESKSNFSCNAL-----KALQKKGIGKESFVCKNGASST 358
Qy 267 KYQFSNYSNGSTGYLPYKLVKAADANNVGLQKLNNGNVQVEFATSTSANNTTANP 326
Db 359 SVKLSSTSKSQSS-----QTTAKVSKSSKAEKKFTSGDIKAAASASSVSSSGASS-- 410
Qy 327 TPAVDEIKVAKIVISGLRFGQNTIELSVPTCEGNNKVPAMIG 369
Db 411 -----SSSKSKGNAIMAPIQO-----TTPLVG 434

RESULT 14
MSPL_PLAFK
ID MSPL_PLAFK STANDARD; PRT; 1630 AA.
AC P04932,
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
DE (PMKSA) (P190).
GN MSP-1.
OS Plasmodium falciparum (isolate K1 / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5839;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86136024; PubMed=3004972;
RA Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,
RA Stunnenberg H., Bujard H.;
RT "Polymorphism of the precursor for the major surface antigens of

RT Plasmodium falciparum merozoites: studies at the genetic level. ";
RL EMBO J. 4:3823-3829(1985).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RA Pan W., Tolle R., Bujard H.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (Potential).
CC -!- PPM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42 kDa AND 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03371; CAA27070.1; --
DR PUR; A25120; SAZOK1
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein; Transmembrane; GPI-anchor.
KW SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1630 MEROZOITE SURFACE PROTEIN 1.
FT DOMAIN 67 84 TRIPEPTIDE SG(TP) REPEAT.
FT TRANSMEM 1614 1630 MEMBRANE ANCHOR.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 755 755 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 759 759 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 774 774 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 835 835 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 955 955 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1630 AA; 187289 MW; ADBDEC3CE0A46322 CRC64;

Query Match 6.3%; Score 126.5; DB 1; Length 1630;
Best Local Similarity 20.7%; Pred. No. 11;
Matches 86; Conservative 62; Mismatches 155; Indels 113; Gaps 19;

Qy 36 LASLDQYAK-----TEASLSAYSEATVNNLNATLQKMAKTNLESAINQANTDKTT 90
Db 226 VKMEDYIKKNNKKTIENTINELIESKKYIDKNKNATKEEK-----KKLYQAYDLSI 278
Qy 91 FDNEHPNLVEAYKALKTTLEQRATNLEGLSSATYNIQIRNLVLDYNKASSLTITKLDPLN 150
Db 279 YNKQ---LEEAHN-LISVLEKRIDTLKK-----NENIKELDKINEI--KNPPPAN 323
Qy 151 GG-----TLDSN-EITTANKNNINNTLSINEOKTNADAL-----SNSFTK 190
Db 324 SGNTPTLTDNKKTEEBEKEIKEAKTI---KFNIDSLFTDPLEYLEYLREKKNKIDIS 380
Qy 191 KVIQNEQSFVGTFTNANVQPSNYSFV-----AFSADVTVPVNYKARTVWNGDEPS 242
Db 381 AKVETKESTEPNEYPNGVTYPLSYNDINNALNELNSFGDLNPFDT-----KGPS 431
Qy 243 SRIAN-----TNSITDVSIIYSLAGTNTKYQFS 271
Db 432 KNIYTDNERKKEINEKIEKKIESDKSYEDRSKSLNDITKEEKL--LNEIYDSK 490
Qy 272 FSNYGPSTGY-----LYFPYKLVKAADANNVGLQKLNNGNVQVEFATSTSANNTTANP 326
Db 491 FNNNIDLTFEKKMGKRYSYKVEKLTH--HTFASYSKSHNLEKLTALKYMEDYSLRN- 548

THIS PAGE BLANK (USPTO)

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	1200	60.1	271	2	Q49464	Q49464 mycoplasma	
2	837	41.9	671	2	Q918D6	Q918D6 mycoplasma	
3	831.5	41.6	702	2	Q49499	Q49499 mycoplasma	
4	826	40.3	632	2	Q9XCGB	Q9XCGB mycoplasma	
5	805	41.3	584	2	Q9KH13	Q9KH13 mycoplasma	
6	761	38.1	680	2	Q9KH14	Q9KH14 mycoplasma	
7	754	37.7	702	2	Q49498	Q49498 mycoplasma	
8	746	37.3	645	2	Q92ID1	Q92ID1 mycoplasma	
9	745	37.3	644	2	Q9ZHR9	Q9ZHR9 mycoplasma	
10	744	37.2	650	2	Q49495	Q49495 mycoplasma	
11	743	37.2	649	2	Q49497	Q49497 mycoplasma	
12	713	35.7	647	2	Q49468	Q49468 mycoplasma	
13	704	35.2	656	2	Q9KH15	Q9KH15 mycoplasma	
14	690.5	34.6	703	2	O05122	O05122 mycoplasma	
15	619.5	31.0	419	2	Q918D5	Q918D5 mycoplasma	
16	575.5	28.8	386	2	Q49500	Q49500 mycoplasma	

Db 147 STAYQIRNNLVLDYNNASSLIKTLDPLNGGMLLDSNEITVNRNINNTLSTINEQKTN 206
QY 181 ADALNSFIKKVIONNEQSFVGTETNANVOPSNYSFVAFSADVTVPVNYKYARRTVWNGDE 240
Db 207 ADALNSFIKKVIONNEQSFVGTETNANVOPSNYSFVAFSADVTVPVNYKYARRTVWNGDE 266
QY 241 PSSRI 245
Db 267 PSSRI 271
RESULT 2
Q9L8D6 PRELIMINARY; PRT; 671 AA.
AC Q9L8D6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE PMGA-LIKE PROTEIN 9.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F;
RA Pharr G.T., Branton S.L., Hanson L.A., Minion F.C., Lott B.D.,
RA May J.B., Hughtlett M.B.; from the F-strain (vaccine strain) of
RT "A novel pmga-like gene from the F-strain (vaccine strain) of
RT Mycoplasma gallisepticum";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF210770; AAF29524.1; -
SQ SEQUENCE 671 AA; 71898 MW; 4D6AE6B59175D679 CRC64;

Query Match 41.9%; Score 837; DB 2; Length 671;
Best Local Similarity 43.4%; Pred. No. 1.1e-33;
Matches 179; Conservative 75; Mismatches 118; Indels 40; Gaps 9;
QY 8 DANP-----NNGQTOLEARMELTDLINAKAMTSLASQDYAKIESLSAYSEAEVTNN 61
Db 55 DTNPGNDGGMENSAQAQAAAKKELSDLLATQNSLNKAYADYTNQNTLTAAYTTAKSTSD 114
QY 62 NLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVAYKALKTTLEQRTATNLEGLSS 121
Db 115 NTSVLEQVKSATSTLQAADTAASAKTSFDEKNPELIRAYALKETLKNETVLSGLTD 174
QY 122 TAYNQIRNNLVLDYNNASSLIKTLDPLNGGMLLDSNEITVNRNINNTLSTINEQKTNA 181
Db 175 SNEATIKTNLTALYQSGKDFVKATLDPVSGNA-PQIADITKADKDIADAVSKLETWKTN 233
QY 182 DALNSFIKKVIONNEQSFVGTETNANVOPSNYSFVAFSADVTVPVNYKYARRTVWNN 237
Db 234 NTLATSFVKEVLKNTLGTIDT-TNNREQPGNYSFVGSYNATNNEIPNWFNAQKRVWT 292
QY 238 GDEPSSRILA-----NTNSITDWSIYSLAGTNTKYQSFNSYGFSTGYLYFPYKLVK 293
Db 293 SDNGRTSLIESTSDNSSTLFEVSWIYSLGAGTKYSLTFNFGPSTGYLYFPYKLVKRGD 352
QY 294 ANNVGLOKLNNGVQVEFA-----TSTSA-----NNTT-----ANPTPAVD 331
Db 353 ENNVALQYTLNLSGSAQEVNAPTQVTSVADSSGDSNNQTESAAETMPVTSDLNPAPTVS 412
QY 332 EIKVAKIVLSLGRGQNTIELSVPTGEGNMKNVAPMIGNIYLSNNENNADKI 383
Db 413 DINIAKTLNLKFGSNTIEFSVPTES--NKVAPMIGNIYLSNNIENAEKV 462
RESULT 3
Q49499 PRELIMINARY; PRT; 702 AA.
ID Q49499
AC Q49499;

DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PMGAI.4 PROTEIN PRECURSOR.
GN PMGAI.4.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum";
RL FEBS Lett. 352:347-352(1994).
DR EMBL: L28424; AAA62418.1; -
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
SQ SEQUENCE 702 AA; 75517 MW; B70AC874FE85055C CRC64;
Query Match 41.6%; Score 831.5; DB 2; Length 702;
Best Local Similarity 42.0%; Pred. No. 2.1e-33;
Matches 173; Conservative 84; Mismatches 116; Indels 39; Gaps 8;
QY 8 DANPNNG-----QTOLEARMELTDLINAKAMTSLASQDYAKIESLSAYSEAEVTNN 61
Db 68 NTPNGGGGTGDNAAQAQAAAKKELSDLLATQNSLNKAYADYTNQNTLTAAYTTAKSTSD 127
QY 62 NLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVAYKALKTTLEQRTATNLEGLSS 121
Db 128 NTSATLEQVKSATSTLQTAIDTAASAKTSFDEKNPELIRAYALKETLKNKRNLSGLTD 187
QY 122 TAYNQIRNNLVLDYNNASSLIKTLDPLNGGMLLDSNEITVNRNINNTLSTINEQKTNA 181
Db 188 SNEATIKTNLTALYQSGKDVTKTLDPLM-GTAINLSAVSQANTNISNAVSKLETWKTN 246
QY 182 DALNSFIKKVIONNEQSFVGTETNANVOPSNYSFVAFSADVTVPVNYKYARRTV 235
Db 247 TVLATSFVKEVLKNTLGTIDT-TNNREQPGNYSFVGSYVDVTTGSDNARPNHSPQKRV 305
QY 236 WNGD-----EPSSRILANTNSITDWSIYSLAGTNTKYQSFNSYGFSTGYLYFPYKLVK 290
Db 306 WTSNTDILSOPQPAEENQOQSAQADVSWIYNLTGMAKYSLTNFGPSTGYLYFPYKLVN 365
QY 291 AADANNVGLQYKLNNGVQVEFATS-----TSANN-----TTANPTPAVD 331
Db 366 SPSDKVALEYKLNESAVKTIQDFSPQTSVPASDATRENNRSTAAAPQAGSTEINPAPTLD 425
QY 332 EIKVAKIVLSLGRGQNTIELSVPTGEGNMKNVAPMIGNIYLSNNENNADK 382
Db 426 DIKIAKTLNLKFGSNTIEFSVPTTAKETGKVPAMIGNIYLSNDRDVK 477
RESULT 4
Q9XCG8 PRELIMINARY; PRT; 632 AA.
AC Q9XCG8;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE VLHAI PRECURSOR (FRAGMENT).
GN VLHAI.
OS Mycoplasma imitans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=29560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4229;

RX MEDLINE=99392472; PubMed=10463176;
RA Markham P.F., Duffy M.F., Glew M.D., Browning G.F.;
RT "A gene family in Mycoplasma imitans closely related to the pmga
family of Mycoplasma gallisepticum.";
RL Microbiology 145:2095-2103(1999).
DR EMBL; AF141940; AAD39483.1; -
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 17 POTENTIAL.
FT CHAIN 18 632 VLHAI.
SQ SEQUENCE 632 AA; 66959 MW; 173F5B12E705BE47 CRC64;

Query Match 41.3%; Score 826; DB 2; Length 632;
Best Local Similarity 45.3%; Pred. No. 3.4e-33;
Matches 183; Conservative 65; Mismatches 120; Indels 36; Gaps 9;

Qy 7 KDANPNNGOT-----OLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSAE 57
Db KPNPDTGNTNPGDGTDAKQADAKASLNTLLGQSTNVVALYEDYAKIKDTLSSAYASQ 93

Qy 58 TVNNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVYAYKALKTTLEORATNLE 117
Db 94 TTANNANATLKDNDKAAQALQTAISDAANAKTEFDKANLGLVSAKALKETLKSETNLD 153

Qy 118 GLSSTAYNOIRNNLDVLYNKASSLITKTLDPLNGGTLTDSNEITTANKNNINTLSTINEQ 177
Db 154 GLSDQNSAIAKANLVSLYNKAKDFTTVLDP-TSGMTPKVDIEITSANTAIQAVSAIDSQ 212

Qy 178 KTNADALNSPIKVIQNEQSPVGTFTN-----ANVQSNYSFVAFSADVTP----- 225
Db 213 KTNADTAATTFEIKELDSAKLT-PGTTAEGQQAASQPGNYSFVGSNDVITGRGSEQ 271

Qy 226 --VNYKYARRVWNGDEPS-----SRILANTNSITDVSWIYSLAGTNTKYQFSNYPSTG 280
Db 272 DLPSWNEAKRKKVWTEGLSAGOTLVSETPLTDVSWIYSLTGAISKYTLFTFYGPSTA 331

Qy 281 YLYFPYKLVKADANNVGLQYKLANG-NVOQVERATPSANNTTANPTPAYDEIKVAKIV 339
Db 332 YLYFPYKLVQSSDNKGLQKLNSETLVITFGNETN-----DSGATPAIDDIKVEKVT 387

Qy 340 LSLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNNENNADKI 383
Db 388 LSNLNFQDNTIEFSVATEE---NKVAPMIGNMYLTSSNNVDKI 428

RESULT 5
Q9KH13 PRELIMINARY; PRT; 584 AA.
AC Q9KH13;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE ADHESIN PMGAI.4 (FRAGMENT).
GN PMGAI.4.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmga multigene family of Mycoplasma
gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR InterPro: IPR001986; EPSP_synase.
DR PROSITE: PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
FT NON_TER 584
SQ SEQUENCE 584 AA; 62453 MW; 5C467BA5F5B72A72 CRC64;

Query Match 40.3%; Score 805; DB 2; Length 584;
Best Local Similarity 42.4%; Pred. No. 3.4e-32;
Matches 180; Conservative 72; Mismatches 113; Indels 60; Gaps 10;

Qy 10 NPNNGOT----OLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATVNNLNA 65
Db 69 NPNSGNTTPEQQLAAARKTLTDLTGENTNVALYADYAKIQSTLSTAYMTAKTASENTSA 128

Qy 56 TLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVYAYKALKTTLEORATNLEGLSSTAYN 125
Db 129 TLDNLSASTTLQAAIDRAASNKRTFDSANQPLVAYNOLKTLTLOSKTTSLEGLSENKYS 188

Qy 126 OIRNNLDVLYNKASSLITKTLDPLNGGTLTDSNEITTANKNNINTLSTNEQKTNADALS 185
Db 189 SIKNHLKSLFDAGSAIAARTLDP-TMGIVPEVMSVTKANEDIMTAVSKLTENKTNADKF- 246

Qy 186 NSFIKKVTQNNQSPVGTFTNAN--VQPSNYSFVAFSADVTP-----VNYKYARTT 234
Db 247 NDFEKKPL--SKEKLVSNDRAHNQEQPANWSFAGYSVDLT*TGSTGNSQNLPNWFAQRK 304

Qy 235 VWNGD---EPSRILANTNSITDVSWIYSLAGTNTKYQFSNYPSTGYLYFPYKLVKA 291
Db 305 VMTSEGOQTGKALVSSPVSATDVSWIYSLAGETKTLSEFYIGPDTAFLIFPYKLVQ 364

Qy 292 ADANNVGLQYKLNNGNVQVEP-----ATST-----S 318
Db 365 ADSSVALQYSLNKTSSKLINEPEAKTPTNADQSENGVATTSTTEGRSSSEVLVADEVA 424

Qy 319 ANNTTANTTPPAYDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNEN 378
Db 425 AVNEMNPTPTVSDINIAKTVLSGLTFGENTIEFSVPT-----NKVAPMIGNMYLTSSNG 479

Qy 379 NADKI 383
Db 480 SOGKI 484

RESULT 6
Q9KH14 PRELIMINARY; PRT; 680 AA.
AC Q9KH14;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE ADHESIN PMGAI.3.
GN PMGAI.3.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmga multigene family of Mycoplasma
gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275312; AAF91414.1; -
SQ SEQUENCE 680 AA; 72909 MW; A2ABC6500E3DFE47 CRC64;

Query Match 38.1%; Score 761; DB 2; Length 680;
Best Local Similarity 39.8%; Pred. No. 5.7e-30;
Matches 164; Conservative 71; Mismatches 129; Indels 48; Gaps 7;

Qy 8 DANPNNGOTQLEAARME-----LTDLINAKAMTSLASQDYAKIEASLSAYSEATVYN 60
Db 63 DTNPGGGGGMNATNQELVNAKKALSDLLGGESKTVELYADYAKIKADLTSAVAKTTS 122

Qy 61 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVYAYKALKTTLEORATNLEGLS 120
Db 123 DSSTSTLDQVKATATSLQTATAINTAASDKKFDQNSQLLMAYKVLKDLTKKKEAIVMSLN 182

121 STAYNQIRNLVDLYNKRASSLIITKTLPLNGGTLDSNEITTANKNINNTL-----ST 173
 183 QEKYSAILSEINAASTAEITVKTLNPVNG-----NLPVVAALNAENTKILFAIKEEK 236
 174 INEOKTNADALNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKYARR 233
 237 INSEKSNADLFANYOLYK-----DRTKLMSEGSNNTKPGNYSFVAYASDIASPNWFAQR 293
 234 TVWNGD-----EPSSRILANTNSITDVSWIYSLAGTNTKYOFSFNSYGPSTGYLYFPYKL 288
 294 TVWTADSTWTSPLPNNLONASPLTDSWIYTLSTGAKYTLITFDYGPQGYLYFPYKL 353
 289 VKAADANNVGLQYKLNGNVQVVEF-----ATSTGANTTANTPTAV 330
 354 VKTSD--KVLQYKLNQADPVAIQSEAAATASAPAEATDGRQESAETATANEKVNPMPSV 411
 331 DEIKVAKIVLSGLRGQNTIELSVPTGEGNMKNKVPMTGNIYLSNENNAK 382
 412 NTINAKVTLSNLKFGSNTIEFSVPMQDDNMKNKVPMTGNIYLSNENNAK 463

RESULT 7
 Q49498 ID Q49498 PRELIMINARY; PRT: 702 AA.
 AC Q49498;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PMGAL.3 PROTEIN PRECURSOR.
 GN PMGAL.3.
 OS Mycoplasma gallisepticum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2096;
 RN [1]
 RC STRAIN=56;
 RX MEDLINE=95010739; PubMed=7925999;
 RA Markham P.F., Grew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
 RA Browning G.F., Whithear K.G., Walker I.D.;
 RT "The organisation of the multigene family which encodes the major cell
 surface protein, pmga, of Mycoplasma gallisepticum."
 RL FEBS Lett. 352:347-352(1994).
 DR EMBL; L28424; AAA62417.1; .
 KW Signal.
 FT SIGNAL.
 SQ SEQUENCE 702 AA; 75537 MW; 273E8915FEE57B9F CRC64;

Query Match 37.7%; Score 754; DB 2; Length 702;
 Best Local Similarity 41.6%; Pred. No. 1.3e-29;
 Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;
 -QY 10 NPNGQT-----QLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATVNNLNA 65
 69 NPNGNTPEQQAAAKRTLLDGTENTVALYADYAKIQSTLSITAYMTAKASENTSA 128
 66 TLEQKMAKTNLESAINQANTDKTTFONEHPNLVEAYKALKTTLEQRTNLEGLSTAYN 125
 129 TLENRSASTTLQAAIDKANDKRVDSVNOPLVAAYNLTKLTKSTTSLEGLSENKYG 188
 126 QIRNVLVDLYNKAASSLITKTLPLNG--GTLLDSNEITTANKNINNTLS--TINEQKTA 181
 189 GIKNHLKSLFDGSAITAKTLTDSRPTLEKYN-----ANNIGKMAISPESLKKWKGN 245
 182 DALNSIFIKKVIQNEQSFVGTFTNANVOPSNYSFVAFSADVTP-----VNYKYARRTV 236
 246 DKF--NEPEKNPLSEKUKLSDTAHQEQPANWFAFAYSDVLTNSQNLPNWNFAQRKVW 304
 237 NGD--EPSSRILANTN--SITDVSWIYSLAGTNTKYOFSFNSYGPSTGYLYFPYKLVAAD 293
 305 TSENQPGKALVSSPVSATDVSWIYSLAGGTGYTLTFFEYIGPDNAFLYLYPKLVKRAAD 364

294 ANNVLQYKLNGNVQVVEF-----ATST-----SAN 320
 365 SSSVALQYSLNKTSSKLNFKPAETVSTNTDQSENEVATTSTTEARSSYKVLVADEAATS 424
 321 NTTANTPTAVDEIKVAKIVLSGLRGQNTIELSVPTGEGNMKNKVPMTGNIYLSNENNA 380
 425 NNNHHTPTVSDINIATKVLTSGLTFGENTIEFSVPEG-----KVAPMIGNMYLTSNSESQ 479
 381 DK1 383
 480 VK1 482

RESULT 8
 Q9ZIDI ID Q9ZIDI PRELIMINARY; PRT: 645 AA.
 AC Q9ZIDI;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
 DE M9 PROTEIN.
 GN M9.
 OS Mycoplasma gallisepticum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2096;
 RN [1]
 RC STRAIN=PG31; ATCC19610;
 RX MEDLINE=99003182; PubMed=9784576;
 RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
 RT "A protein (M9) associated with monoclonal antibody-mediated
 agglutination of Mycoplasma gallisepticum is a member of the pmga
 family."
 RT Infect. Immun. 66:5570-5575(1998).
 DR EMBL; AF032890; AAC69269.1; .
 SQ SEQUENCE 645 AA; 69876 MW; 5A2F29572CFAB055 CRC64;

Query Match 37.3%; Score 746; DB 2; Length 645;
 Best Local Similarity 42.3%; Pred. No. 2.9e-29;
 Matches 172; Conservative 64; Mismatches 123; Indels 48; Gaps 11;
 QY 8 DANPNNGQ-----TOLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSEATVN 60
 44 DTNPGDQGGMMNAAOELAAARMGLITVFDKAKNLGLYVDYKKTQDTLTKAYDAKTVL 103
 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 120
 104 DNSSTTQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKVYELKTTLSNETATLAPYA 163
 121 STAYNQIRNLVDLYNKAASSLITKTLPLNGGTLDSNEITTANKNINNTL--STINEOK 178
 164 AAOYAGIKMHLGSLYDAGAKITTKTLEPEVGGDP--LTADVVMMANTKIVEAIKDEVLPQK 222
 179 TNADALNSIFIKKVIQNEQSFVGTFTNANVOPSNYSFVAFSADVTPV-----NRYK 230
 223 ENATKADSFVKVLYKEKITGVEEAHN--RAQPANYSFVGVSDITGTVTQTSIPNWDY 281
 231 ARRTVW--NGDEPPSRILANT-----NSITDVSWIYSLAGTNTKYOFSFNSYGPSTGY 281
 282 AORTIFTNGDEP--RSISNTPADQGTWVQPLSNYSWIYSLAGTGAKTLEFTFYGPSTGY 339
 282 LYFPYKLVKAADANNVGLQYKLNGNVQVVEFATSTSA-----NNTANPTPAYDEIKVA 336
 340 LYFPYKLVNTSDQMKLGLEYKLND-----ATEPSAITGSECTMNGKTKPTVNDINVA 391
 337 KIVLSGLRGQNTIELSVPTGEGNMKNKVPMTGNIYLSNENNAKDI 383
 392 KVTLANLKFGSNKIEFSVPA-----EKVSPMIGNMYLSSSPNNNNKI 433

RESULT 9

```
Q92HR9
ID Q92HR9 PRELIMINARY; PRT; 644 AA.
AC Q92HR9;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)
DE HYPOTHETICAL 69.9 KDA PROTEIN.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PG31, ATCC19610;
RC MEDLINE-99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RA "A protein (M9) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmga
RT family."
RL Infect. Immun. 66:5570-5575(1998).
DR EMBL: AF053978; AAC69274.1; -.
KW Hypothetical protein.
SQ SEQUENCE 644 AA; 69866 MW; 8B9F352B13FDE5C CRC64;

Query Match 37.3%; Score 745; DB 2; Length 644;
Best Local Similarity 42.3%; Pred. No. 3.3e-29;
Matches 172; Conservative 65; Mismatches 122; Indels 48; Gaps 11;

Qy 8 DANPNQO-----TQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTN 60
Db 44 DTPGDGGMNAAQELAAARMGLTTVFDSKAKNLGLYVDYKKTQDTLTAKYDAAKTVL 103
Qy 61 NNLNATLEOLKMAKTNLESAINQANTDKTFDNEHPNLVEAYKALKITLLEORATNLEGLS 120
Db 104 DNSSTQNLNEAKRLETAIRTAATSKQTFDEQHAELVKYVEELKTLTSLNETATLAPYA 163
Qy 121 STAYNQIRNNLDVLYNKASSLITKTLPLNGTLLDSNEITANKNNNTL--STINBQK 178
Db 164 AAQYAGIKMHLGSLYDAGKAITTKTLEPVEGDP-LTASAVMMANTKIVEAIKDEVLPQK 222
Qy 179 TNADALNSFTKKYQIONNEQSFVGTFTNANVPQSNYSFVAFSADVTPV-----NKKY 230
Db 223 ENATKLADSFVKQVLVKEKITGVEEAHKS-QPANYSFVGSVDITGTANGOTSIPNNNY 281
Qy 231 ARRTVW-NGDEPSSRLANT-----NSITDYSWIYSLAGTNTKYQFSNYPSTGY 281
Db 282 AQRFTNGDEP--RSVSNTFPVGDQTMQAQPLSNVSWIYSLAGTCAKYTLFTYYPSTGY 339
Qy 282 LYFPYKLVKAADANNVGLQYKLNNGVQVVEFATSTSA-----NNTTANPTPAYDEIKVA 336
Db 340 LYFPYKLVNTSDQVKGLEYKLN-----ATKPSAITGSEQTMNGKTPVNDINVA 391
Qy 337 KIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKI 383
Db 392 KVTLANLFGSNKIEFSVP-----VEKVPIMIGNIYLSNENNADKI 433

RESULT 10
Q49495 PRELIMINARY; PRT; 650 AA.
AC Q49495; O08060;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DE 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE HAEMAGGLUTININ PRECURSOR.
GN PMGAL.1.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.

Q49497 PRELIMINARY; PRT; 649 AA.
AC Q49497;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PMGAL.2 PROTEIN PRECURSOR.
GN PMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S6;
RC MEDLINE-95010739; PubMed=7925999;
RA Markham P.F., Grew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RA "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum."
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S6;
RC Markham P.F.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U0714; AAB50152.1; -.
KW SIGNAL.
ET CHAIN 1 25 POTENTIAL.
ET CHAIN 26 650 HAEMAGGLUTININ.
SQ SEQUENCE 650 AA; 70249 MW; 3ABACDB65940EBBB CRC64;

Query Match 37.2%; Score 744; DB 2; Length 650;
Best Local Similarity 42.9%; Pred. No. 3.7e-29;
Matches 172; Conservative 62; Mismatches 127; Indels 40; Gaps 11;

Qy 10 NPNNGQ-----TQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVTNNN 62
Db 51 NPGDQGMNAAQELAAARMGLTTVFDSKAKNLGLYVDYKKTQNTLTAKYDAAKTVLDN 110
Qy 63 LNATLEOLKMAKTNLESAINQANTDKTFDNEHPNLVEAYKALKITLLEORATNLEGLST 122
Db 111 SSSTQNLNEAKRLETAIRTAATSKQTFDEQHAELVKYVEELKTLTSLNETATLAPYAD 170
Qy 123 AYNQIRNNLDVLYNKASSLITKTLPLNGTLLDSNEITANKNNNTL--STINBQKN 180
Db 171 QYAGIKMHLGSLYDAGKAITTKTLEPVEGDP-LTACAVTMANTKIVEAIKDEVLPNPKKN 229
Qy 181 ADALNSFTKKYQIONNEQSFVGTFTNANVPQSNYSFVAFSADVTPV-----NKKYAR 232
Db 230 ATKLADSFVKQVLVKEKITGVEEAHN-KAOPANYSFVGSVDITGTANGOTSIPNNNYAQ 288
Qy 233 RTVW-NGDEPSSRLANT-----NSITDYSWIYSLAGTNTKYQFSNYPSTGYLY 283
Db 289 RTFTNGDEP--RSVSNTFPVGDQTMQAQPLSNVSWIYSLAGTCAKYTLFTYYPSTGYLY 346
Qy 284 FPKLVKAADANNVGLQYKLNNGVQVVEFATSTANNTAN--PTPAVDEIKVAKTVLGS 342
Db 347 FPKLVNTSDQVKGLEYKLNDA---TEPSAITGNEQTMNGKTPVNDINVAKYTLAN 402
Qy 343 LRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKI 383
Db 403 LIFGSKNIEFSVPA-----EKVSPMIGNIYLSNENNADKI 438

RESULT 11
Q49497 PRELIMINARY; PRT; 649 AA.
ID Q49497;
AC Q49497;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PMGAL.2 PROTEIN PRECURSOR.
GN PMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S6;
RC MEDLINE-95010739; PubMed=7925999;
RA Markham P.F., Grew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RA "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum."
RN [2]
RP SEQUENCE FROM N.A.
```

RL FEBS Lett. 352:347-352(1994).
DR EMBL; L28424; AAA62416.1; -
KW Signal.
FT SIGNAL 1 25
* SO SEQUENCE 649 AA; 70205 MW

Query Match 37.2%; Score 743; DB 2; Length 649;
Best Local Similarity 42.3%; Pred. No. 4.1e-29;
Matches 172; Conservative 62; Mismatches 125; Indels 48; Gaps 11;

Qy	8	DANPNNGQ-----TQLEAARNELTDLINAKAMTILASLODYAKTEIASLSSYASAEATVN	60
Db	48	DTNPGDQGMNNAASQELAAARGLTTTFDSRAKUNGLYVDYKKTQNTLTRAYDAAKTVL	107
Qy	61	NNLNAYLEQLKMAKTUNLESAINQANTDKTTTFONEHPNLVEAYKALKTTLEQRATNLEGLS	120
Db	108	DNSSTTONLEAKTRLETAIRTAATSKOTFEQHAELKVYKELKTTLSNETATLAPYA	167
Qy	121	STAYNOIRNNLVLYNKASSLITKTLDPDLNGTGLDSDNEITTANKNINWTL--STINEQK	178
Db	168	DAQYAGIKMHLGLYDAGKAITTKTLEPVEGDP-LTASAVMANTKIVEAIKDEVLNPOK	226
Qy	179	TNADALSNSFIKKVYQIONNEOSFGVGTTNANQPSNYSFVAFSADTVPV-----NYKY	230
Db	227	ENATKLADSEVKQVLKVEKITGKITEEAAHN-KAQPANPSFGVSDIITGTTGQTSIPNWDY	285
Qy	231	ARRTVW--NGDEPSSRIILANT-----NSITDVSMTYSLAGTNTTKYQFSFSGNYGSTCY	281
Db	286	AQRTFTNSDEP--RSLSNTPADQOTWAQPLSNVSMIYSLAGTCAKYLTFTYVYGSTCY	343
Qy	282	LYFFYKLVKAADANNVGLQKLANGNYQVEFATSTSA-----NNTTANPTPAVDKIVA	336
Db	344	LYFFYKLVNTSDQVKLGLEYKLN-----ATKPSAITGSDQTMNGKTPTVNDINVA	395
Qy	337	KIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSSENNAADKI	393
Db	396	KVTLANTFGNSKTFEFSVPA-----EKVSPMIGNIYLSSENPNWNKI	437

RESULT	12	
Q49468		
ID	Q49468	PRELIMINARY; PRT; 647 AA.
AC	Q49468: Q53303;	
DT	01-NOV-1996	(trEMBLrel. 01, Created)
DT	01-NOV-1996	(trEMBLrel. 01, Last sequence update)
DT	01-MAY-2000	(trEMBLrel. 13, Last annotation)
DE	HEMAGGLUTININ HOMOLOG PRECURSOR.	
GN	PMGAL.2	
OS	Mycoplasma gallisepticum.	
OC	Bacteria; Firmicutes; Bacillus/Clostridium	
OC	Mycoplasmataceae; Mycoplasma.	
OX	NCBI_TaxID=2096;	

AN	SEQUENCE FROM N.A.		
RP	MEDLINE=93162830; PubMed=8432610;		
RX	Markham P.F., Glew M.D., Whithear K.G., Walker I.D.;		
RA	"Molecular cloning of a member of the gene family that encodes pmGA, a		
RT	hemagglutinin of Mycoplasma gallisepticum.";		
RT	Infect. Immun. 61:903-909(1993).		
RL	EMBL; M83178; AAA02996.1; -		
DR	EMBL; S55216; AAB25397.2; -		
DR	Signal.		
KW	POTENTIAL.	1	25
FT	SIGNAL.	26	647
FT	CHAIN		HEMAGGLUTININ HOMOLOG.
SO	SEQUENCE 647 AA; 70333 MW; 33916673BB9E28C4 CRC64;		

Query Match	35.7%	Score 713;	DB 2;	Length 647;
Best Local Similarity	41.1%	Pred. No. 1.2e-27;		
Matches 169.	Conservative	60;	Mismatches 124;	Indels 58;
				Gaps 12;

0V 8 DANPNNGQ-----TQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAE TVN 60

Db	48	DTNPGDGGOMNAASQELAAARMGLTTITFDSKAKNGLYVDYKKQTNTLTKAYDAAKTVL	107
Qy	61	NNLNATLEOLKMAKTNLESAINQANTDKTTFONEHPNLVYAYKALKTTLLEQATNLEGLS	120
Db	108	DNSSSTONLNEAKTRLEYTAIRTAATSQTDFDSQHAELVKVYKELKTTLSNETATLAPYA	167
Qy	121	STAYNQIRNNLVLYNKASSLITKTLTDLPLNGTLLDSNEITANKNINWTL--STINEOK	178
Db	168	DAQYAGIKMHLGSLYDAGKAITTKTLEPVEGDP-LTASAVMAMANTKIVEAIDKLVNPK	226
Qy	179	TN----ADALNSFIKVIQNNQSFGTGTNANVOPSNYSVFASADVTPV-----	226
Db	227	ENATKLAADSLLSLVKKITGVEE-----AHNKAOPANYSFVGYKRWTELLLDKQVEP	279
Qy	227	NYKVARVTW--NGDPPSSRLANT-----NSITDVSWIYSLAGTNTKYQFSFSNYGP	277
Db	280	NWDYAQRTFTNSDEP--RSISNTPADQOTMAOPLSNVSWIYSLAGTGAKYTLFEITYGP	337
Qy	278	STGYLYPPYKLYKAADANNVGLQKLANGNVQOVEPATSTSA-----NNTTAMPTPAVDE	332
Db	338	STGYLYPPYKLVNTSDQVKLGLEYKLD-----ATKPSAITFGSDQTMNGKTPTVND	389
Qy	333	IKVAKIVLSGLRFQONTIELSVPTGEGNMKNVPMIGNIYLSNENNADKI	383
Db	350	TNVAQVTIANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSPNWNKKI	435

RESULT 13
OQKH15

Q9KH13	PRT: 656 AA.
TD Q9KH15	PRT: 656 AA.

Q9KRH5;	FALLINGWATER,
AC	01-OCT-2000 (TREMBlrel. 15, Created)
DT	01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT	01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DT	01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE	ADHESIN PMGAL.2.
DE	PMGAL.2.
GN	PMGAL.2.
OS	Mycoplasma gallisepticum.
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC	Mycoplasmataceae; Mycoplasma.
OX	NCBI_TaxID=2096;
OX	[1]
RN	SEQUENCE FROM N.A.
RP	STRAIN=HS;
RC	Shen Q.C., Bi D.R., Weng C.J.;
RA	"Sequence analysis of the pmCA multigene family of Mycoplasma
RT	gallisepticum strain HS."
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RL	EMBL; AF275312; AAF91413.1; "
DR	SEQUENCE 656 AA; 70875 MW; 58EA7E075FC617E1 CRC64;
SO	

Query Match	35.2%	Score 704;	DB 2;	Length 656;
Best Local Similarity	41.0%	Pred. No. 3.4e-27;		
57. Mismatches	124:	Indels	48:	Gaps
12:				

Qy	8	DAMPNGQT-----QLEARMELTDLLNAXAMTLLASQDQKATLASLSSAYSAEATVN	60
Db	63	DINPGGQNMDSAAQELTAARTALTSLLSKANANVEMISDYAKIQTIIIAAYTAEQTS	122
Qy	61	NNLNATLEQLKMAKTNLUESAINOANTDKTTFONEHPNLVYAEYKAKTTLLEQRATNLLEGUS	120
Db	123	QNSSATLEQVKNTATSLQTAINTANSNKQKFDQDHSNLLWSYKNLMATLAKKETA VMTLK	182
Qy	121	STAYNQIRNNLVDILNYKASSLTITKLDPLNGGTLSDSEITITANKNINWTLs--TINEOK	178
Db	183	DPKYSALDQINGVSSKGEELVQHTLDPVS-GIVPAANTITEITIKIEEVISEKTLQDQK	241
Qy	179	TNADALS--SFI--KKVIQNNQOSFVGTTTNAVOPSNYSFVAFSADYTPV-----	226
Db	242	NNADQAFYQSFTLDTKTLNVEDA-----KKMGOPANYSFVGYSVDVTGTSGQETTP	295
Qy	278	-----ITDYSWITSYSLAGTNTKYOFSSFNYSNGPS	278

```

Db 296 MNNEAQAIFTSNGNPTKVATTTGDDQSTAKPLSDVSWIYSLAGTAKYTLTEFTYIGPS 355
Qy 279 TGCLYFPYKLVKAADANNVGLQYKLN-NGNQVQVFATSTSANNTAPTAPVDEIKVAK 337
Db 356 TGWLYFPYKLVKAND--DVGLOQYKLSNETLPIIFGGT-----TTNGPAATVENINAK 409
Qy 338 IVLSGLRFGONTIELSVPTGEGNMKNKVPAMIGNIYLSNNENNAK 382
Db 410 VRLTGLAFGKNTIEFVSP-----MSKVAPMIGNMYITSSDTEITNK 449

RESULT 14
ID O05122 PRELIMINARY; PRT; 703 AA.
AC O05122;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HAEMAGGLUTININ.
GN PMGAL.9.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F.; Glew M.D.; Sykes J.E.; Bowden T.R.; Pollocks T.D.;
RA Browning G.F.; Whithear K.G.; Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmGA, of Mycoplasma gallisepticum.";
RL FEBS Lett. 352:347-352(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RA Markham P.F.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90714; AAB50154.1; -.
DR InferPro; IP002819; HD.
SQ SEQUENCE 703 AA; 75742 MW; 310B69BE9F73CB5 CRC64;

Query Match 34.6%; Score 690.5; DB 2; Length 703;
Best Local Similarity 35.0%; Pred. No. 1.7e-26;
Matches 167; Conservative 75; Mismatches 120; Indels 115; Gaps 12;

Qy 1 CMSTTKDANP-----NNGQT-----OLEAARMELTDLINAK 32
Db 26 CTSATIPTLNPPEKPPDPMPNPPSGMNGGNTNPGMDTAQELASAKAALTITNRE 85
Qy 33 AMTLASQDYAKIEASLSAYSEATVNNLNATLEQLKMAKTNLESAINOANTDKTTFD 92
Db 86 SEKVGLYDYAKIKADLSAYTAVKTTSDSSTSLVQVKTATSKLQTAIDKAASQKQFE 145
Qy 93 NEHPNLVEAYKALKTLEQR-ATNLEGLSSTAYNOIRNLVDLYNKASLITKTLDPNG 151
Db 146 QHKDLLMPYSELKTTLSQKNATVL--LNQPKYSAILNKINSIYAQGEVWIRLTDPSV- 202
Qy 152 GTLDSNEITTANKNINNTLS--TINEOKTNADALSNSFIKKVIONNEQSFVGTFTNANY 209
Db 203 GAIPTAASITKVNDEINKAISENQLKPKKNADAFANYQFFKL---DKTKINGMSTNMK 259
Qy 210 QPSNYSFVAFSADYTPV-----NYKARTVWNGDEPSSRILANTNS----- 251
Db 260 QPQNTYFVGYSGVGTGMSGQTTIPNWNFAORIVWSSGAPRAPLASQTETPOAETPPMSA 319
Qy 252 -----ITDVSWIYSLA 262
Db 320 POGVEPAQQGSDSPKQASEQVSEPTFAAEVQAQADTEQPATSGQTPLTDDVSWIYSL 379
Qy 263 GTNTKYQFSFSGYSGTGYLYFPYKLVKAADANNVGLQYKLNNGNVQVFEATSTSANNT 322
```

```

Db 380 GTDVKYTFTEFNFGPSMAYLYFPYKLVKSD--SVGLQYKLNNNNPVALNFGSETNAN-- 435
Qy 323 TANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNMKNKVPAMIGNIYLSNNEN 379
Db 436 --GPAASVDNINIVAKVNLNLANLNFGEITIEFVSP-----MNKVAPMIGNMYITSDVAN 485

RESULT 15
ID Q9L8D5 PRELIMINARY; PRT; 419 AA.
AC Q9L8D5;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PMGA-LIKE PROTEIN 9.3 (FRAGMENT).
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F;
RA Pharr G.T.; Branton S.L.; Hanson L.A.; Minion F.C.; Lott B.D.;
RA May J.D.; Hughlett M.B.;
RT "A novel pmGA-like gene from the F-strain (vaccine strain) of
RT Mycoplasma gallisepticum.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF210770; AAF29525.1; -.
FT NON_TER 419
SQ SEQUENCE 419 AA; 45072 MW; FE5EBE37F2DB3B0C CRC64;

Query Match 31.0%; Score 619.5; DB 2; Length 419;
Best Local Similarity 40.1%; Pred. No. 2.9e-23;
Matches 149; Conservative 58; Mismatches 122; Indels 43; Gaps 11;

Qy 8 DAPNNGQT-----OLEAARMELTDLINAKAMTDLASLDQYAKIEASLSAYSEATVN 60
Db 61 DINPGGQNMDSAAQELTAARTALTSLLSKNANTEMYSYAKIQTWLIATTAQT 120
Qy 61 NNLNATLEQLKMAKTNLESAINOANTDKTTFDNEHPNLVEAYKALKTTLEQRAINLEGLS 120
Db 121 QNSATLEQVKNATSALQTAINTANSNKQKQDQDHSNLLMSYKNLMATLAKKETTMTLK 180
Qy 121 STAYNOIRNLVDLYNKASLITKTLDPNGGLTDSNEITTANKNINNTLS--TINEOK 178
Db 181 DPKYSAILDQINGVSCGEEVQHTLDPYS-GIVPAANTITEEITKIEEVISEKTLQDQK 239
Qy 179 TNADALSN--SFI--KKVIONNEQSFVGTFTNANVOPSNYSFVAFSADYTPV----- 226
Db 240 NNADQFNTQSTFLDKTKLUNVEDA-----KKMGQPNYSFVGVSDVTGSGQETTP 293
Qy 227 NYKARTVWNGDEPSSRILANTNS-----ITDVSWIYSLAGTNTKYQFSFSGY 278
Db 294 NNFEAQAIFTSNGNPTKVATTTGDDQSTAKPLSDVSWIYSLAGTAKYTLTEFTYIGPS 353
Qy 279 TGCLYFPYKLVKAADANNVGLQYKLN-NGNQVQVFATSTSANNTAPTAPVDEIKVAK 337
Db 354 TGWLYFPYKLVKAND--DVGLOQYKLSNETLPIIFGGT-----TTNGPAATVENINAK 407
Qy 338 IVLSGLRFGONT 349
Db 408 VRLTGLAFGKNT 419

Search completed: June 12, 2002, 10:50:40
Job time: 203 sec
```

us-09-147-052-2_copy_64_456.rspt

Thu Jun 13 10:00:48 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:48:38 ; Search time 55.68 Seconds

(without alignments)
785.975 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086

Perfect score: 2004

Sequence: 1 GCMSTKKDANPNNGTQLE.....SSNENADKIPGRRPGTEL 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_032802.*
- 1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
 - 2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
 - 3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
 - 4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
 - 5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
 - 6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
 - 7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
 - 8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
 - 9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
 - 10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
 - 11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
 - 12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
 - 13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
 - 14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
 - 15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
 - 16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
 - 17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
 - 18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
 - 19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
 - 20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
 - 21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
 - 22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2004	100.0	456	18	AAW36050 Hybrid Marek's dis
2	2000	99.8	1086	18	AAW36051 Hybrid Marek's dis
3	1910	95.3	615	15	AAW63230 Mycoplasma gallise
4	1856	92.6	610	15	AAW63229 Mycoplasma gallise
5	1612	80.4	368	14	AAW44493 Mycoplasma gallise
6	1612	80.4	368	15	AAW63227 Mycoplasma gallise
7	1604	80.0	368	16	AAW76955 Mycoplasma gallise
8	1142	57.0	235	10	AAW93646 Amino acid sequenc
9	1142	57.0	235	11	AAW05081 MG-1 antigen. AA
10	1142	57.0	261	11	AAW05082 TMG-1 antigen. A
11	1142	57.0	261	16	AAW79911 M.gallisepticum 26

12	1117	55.7	261	10	AAW93959 Amino acid (AA) se
13	806	40.2	661	15	AAW63226 Mycoplasma gallise
14	806	40.2	661	16	AAW79910 M.gallisepticum 66
15	713	35.6	647	15	AAW11978 Mycobacterium gall
16	700.5	35.0	648	15	AAW56973 PMGA 1.2 protein o
17	386.5	19.3	183	10	AAW93649 Amino acid sequenc
18	386.5	19.3	183	11	AAW06439 MG-4 antigen. AA
19	284.5	14.2	219	10	AAW93648 Amino acid sequenc
20	284.5	14.2	219	11	AAW06438 MG-3 antigen. AA
21	176.5	8.8	6281	22	AAW37403 Staphylococcus aur
22	176	8.8	1095	22	AAW83030 S. epidermidis ope
23	171	8.5	2086	22	AAW34143 Staphylococcus aur
24	171	8.5	5795	22	AAW37017 Staphylococcus aur
25	162	8.1	1029	22	AAW34389 Staphylococcus aur
26	162	8.1	1048	22	AAW37490 Staphylococcus aur
27	159	7.9	933	21	AAW58435 Staphylococcus aur
28	159	7.9	933	22	AAW89508 Staphylococcus aur
29	159	7.9	936	18	AAW89801 Staphylococcus aur
30	158	7.9	2434	22	AAW34339 Staphylococcus aur
31	156	7.8	807	21	AAW18311 Plasmodium falcipa
32	154.5	7.7	682	17	AAW95273 Nisin nisp gene pr
33	153.5	7.7	2478	22	AAW34320 Staphylococcus aur
34	153.5	7.7	2478	22	AAW37374 Staphylococcus aur
35	151.5	7.6	1237	18	AAW55640 H. pylori ORF 04ep
36	151.5	7.6	1237	20	AAW17187 H. pylori outer me
37	151	7.5	5024	22	AAW82935 S. epidermidis ope
38	145.5	7.3	2437	22	AAW34338 Staphylococcus aur
39	145	7.2	1215	22	AAW34412 Staphylococcus aur
40	145	7.2	1269	22	AAW37520 Staphylococcus aur
41	143.5	7.2	837	22	AAW34387 Staphylococcus aur
42	143.5	7.2	875	22	AAW37487 Staphylococcus aur
43	143	7.1	3158	22	AAW37018 Staphylococcus aur
44	142	7.1	135	15	AAW56974 Partial PMGA 1.3 p
45	142	7.1	135	15	AAW11979 Mycobacterium gall

ALIGNMENTS

RESULT 1

AAW36050

ID AAW36050 standard; Protein; 456 AA.

XX AC AAW36050;

XX DT 15-JUL-1998 (first entry)

XX DE Hybrid Marek's disease virus/M. gallisepticum 40 K-S protein.

XX KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein;

XX KW antigen; vaccine; poultry.

XX OS Chimeric - Marek's disease gammaherpesvirus.

XX OS Chimeric - Mycoplasma gallisepticum.

XX FH Key Location/Qualifiers

XX FT Region 1..64

XX FT Region /note= "derived from Marek's disease virus gB protein"

XX FT Region 65..456

XX FT Region /note= "derived from M. gallisepticum antigenic protein"

XX PN WO9736924-A1.

XX PD 09-OCT-1997.

XX PF 28-MAR-1997; 97WO-JP01084.

XX PR 29-MAR-1996; 96JP-0103548.

XX PA (JAPG) NIPPON ZEON KK.

XX PI Saito S, Tsuzaki Y, Yanagida N;

XX XX

DR WPI: 1997-503046/46.
DR N-PSDB; AAT96595.
XX
XX Fusion protein comprising herpes virus outer membrane protein and
PT antigenic polypeptide - for prevention of infection by Mycoplasma
PT gallisepticum, especially in poultry
XX
XX PS Disclosure: Page 16-19; 51pp; Japanese.
XX
XX This sequence represents the chimeric protein 40 K-C which comprises a
CC fragment of the Marek's disease virus outer membrane protein gB fused
CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
CC protein can be used in recombinant live vaccines for prevention of
CC infection by Mycoplasma gallisepticum, especially as the outer membrane
CC protein shows antigenicity in poultry.
XX
XX SQ Sequence 456 AA;

Query Match 100.0%; Score 2004; DB 18; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.3e-128;
Matches 394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCMSTTKDANPNNGOTLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSEAEV 60
DB 63 GCMSTTKDANPNNGOTLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSEAEV 122
QY 61 NNNLNLATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 120
DB 123 NNNLNLATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 182
QY 121 SSTAYNQIRNNLDLYNKASSLITKTLDPINGTLLDSNEITTANKNINNTLSTINEQKT 180
DB 183 SSTAYNQIRNNLDLYNKASSLITKTLDPINGTLLDSNEITTANKNINNTLSTINEQKT 242
QY 181 NADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTWNGD 240
DB 243 NADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTWNGD 302
QY 241 EPSSRLANTNSITDVSWIYSLAGTNKYQFSFNSYGPSTGYLYFPYKLVKAADANNVGL 300
DB 303 EPSSRLANTNSITDVSWIYSLAGTNKYQFSFNSYGPSTGYLYFPYKLVKAADANNVGL 362
QY 301 OYKLNNQNVQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEG 360
DB 363 QYKLNNQNVQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEG 422
QY 361 NMKVAPMIGNIYLSNNENNADKIPGRRPGTFL 394
DB 423 NMKVAPMIGNIYLSNNENNADKIPGRRPGTFL 456

RESULT 2
ID AAW36051 standard; Protein; 1086 AA.
XX
XX AAW36051;
XX
XX 15-JUL-1998 (first entry)
XX
XX Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.
XX
XX Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
KW antigen; vaccine; poultry.
XX
XX Chimeric - Marek's disease gammaherpesvirus.
OS Chimeric - Mycoplasma gallisepticum.
XX
XX Key Location/Qualifiers
FT Region 1..672 "derived from Marek's disease virus gB protein"
FT /note= 693..1086
FT /note= "derived from M. gallisepticum antigen"

XX WO9736924-A1.
XX
XX PD 09-OCT-1997.
XX
XX PF 28-MAR-1997; 97WO-JP01084.
XX
XX PR 29-MAR-1996; 96JP-0103548.
XX
XX PA (JAPG) NIPPON ZEON KK.
XX
XX PI Saito S, Tsuzaki Y, Yanagida N;
XX
XX DR WPI: 1997-503046/46.
XX
XX DR N-PSDB; AAT96596.
XX
XX PT Fusion protein comprising herpes virus outer membrane protein and
PT antigenic polypeptide - for prevention of infection by Mycoplasma
PT gallisepticum, especially in poultry
XX
XX PS Disclosure: Page 22-30; 51pp; Japanese.
XX
XX This sequence represents the chimeric protein 40 K-C which comprises a
CC fragment of the Marek's disease virus outer membrane protein gB fused
CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric
CC protein can be used in recombinant live vaccines for prevention of
CC infection by Mycoplasma gallisepticum, especially as the outer membrane
CC protein shows antigenicity in poultry.
XX
XX SQ Sequence 1086 AA;

Query Match 99.8%; Score 2000; DB 18; Length 1086;
Best Local Similarity 99.7%; Pred. No. 7.5e-128;
Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCMSTTKDANPNNGOTLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSEAEV 60
DB 693 GCMSTTKDANPNNGOTLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSEAEV 752
QY 61 NNNLNLATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 120
DB 753 NNNLNLATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 812
QY 121 SSTAYNQIRNNLDLYNKASSLITKTLDPINGTLLDSNEITTANKNINNTLSTINEQKT 180
DB 813 SSTAYNQIRNNLDLYNKASSLITKTLDPINGTLLDSNEITTANKNINNTLSTINEQKT 872
QY 181 NADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTWNGD 240
DB 873 NADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTWNGD 932
QY 241 EPSSRLANTNSITDVSWIYSLAGTNKYQFSFNSYGPSTGYLYFPYKLVKAADANNVGL 300
DB 933 EPSSRLANTNSITDVSWIYSLAGTNKYQFSFNSYGPSTGYLYFPYKLVKAADANNVGL 992
QY 301 OYKLNNQNVQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEG 360
DB 993 QYKLNNQNVQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEG 1052
QY 361 NMKVAPMIGNIYLSNNENNADKIPGRRPGTFL 394
DB 1053 NMKVAPMIGNIYLSNNENNADKIPGRRPGTFL 1086

RESULT 3
ID AAW63230 standard; Protein; 615 AA.
XX
XX AAW63230;
XX
XX 23-JUN-1995 (first entry)
XX

QY 242 PSSRLANTNSITDVSWIYSLAGTNTKQFSNSVGPSTGHLVPPYKLVKAADANNVGLQ 301
 |||||
 DB 267 pssrlantnsitdvswiyslsgtntkqfsnysgpgstgylyfpyklvkaadasnvgllq 326
 |||||
 QY 302 YKLNGNVQOQVEFATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGSGN 361
 |||||
 DB 327 yklngnvpqvefatstsannttanptpavdeikvakivlsglrfgqntielsvptgern 386
 |||||
 QY 362 MNKVAPMIGNIYLSNENNAADK 383
 |||||
 DB 387 mnkvapmignmyitssnaeank 408
 |||||
 RESULT 5
 AAR44493
 ID AAR44493 standard; Protein; 368 AA.
 XX AAR44493;
 AC AAR44493;
 DT 16-JUN-1994 (first entry)
 XX
 DE Mycoplasma gallisepticum 40kd antigen.
 XX
 KW Vaccine; mycoplasma infection; poultry; fowl.
 XX
 OS Mycoplasma gallisepticum.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 262 /note= "corresponds to NNN codon in AAR44493"
 FT Misc-difference 283 /note= "corresponds to NNN codon in AAR44493"
 FT
 XX W09324646-A.
 PN
 XX
 PD 09-DEC-1993.
 XX
 PF 28-MAY-1993; 93WO-JP00715.
 XX
 PR 29-MAY-1992; 92JP-0138819.
 XX
 PA (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX
 PI Aoyama S, Fujisawa A, Iritani Y, Ohkawa S, Saito S;
 XX
 DR WPI; 1993-405837/50.
 DR N-PSDB; AAR44493.
 XX
 PT Mycoplasma gallisepticum antigen and DNA coding for it - useful
 PT for vaccination of fowl against mycoplasma infections
 XX
 PS Claim 2; Page 23-26; 37pp; Japanese.
 -XX
 CC The sequence coding for the 40kDa antigen was obtained by PCR
 CC amplification of M.gallisepticum genomic DNA. The antigen reacts
 CC with Mycoplasma-immune or Mycoplasma-infected serum and can be used
 CC as a vaccine to protect fowl from M.gallisepticum infection.
 XX
 SQ Sequence 368 AA;

Query Match 80.4%; Score 1612; DB 14; Length 368;
 Best Local Similarity 95.0%; Pred. No. 4.8e-102;
 Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNQTOLEARMELDTLINAKANTLASLODYAKIEASLSAYSEATVN 61
 |||||
 DB 27 cmsitkdkanpnngqtdqarmeltdlinakartlasldyakielaslsayseactn 86
 |||||
 QY 62 NNLNATLEQLKMAKTINLESAINQANTDKTTFDNEHPNVLVEAYKALKTTLEQRATNLEGLS 121
 |||||

DB 87 mnlntleqlkmaktnlesainqantdkttfdnehpnlveaykalkttleqratnlegla 146
 QY 122 STAYQIQRNNLVLYNKASSLITKTLDPNGCTLLDSNEITTANKNNNTLSTINBQKTN 181
 |||||
 DB 147 stayqirnnlvlynnasslitktdlpinggmildsneittvnnntlnntslcineqkcn 206
 |||||
 QY 182 ADALNSFIKKVIONEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDE 241
 |||||
 DB 207 adalnsfikkviqneqsfvgtftnanvqpsnysfafsadvtpvnykyarrtvxngde 266
 |||||
 QY 242 PSSRLANTNSITDVSWIYSLAGTNTKQFSNSVGPSTGHLVPPYKLVKAADANNVGLQ 301
 |||||
 DB 267 pssrlantnsitdvsxiyslagtntkqfsnysgpgstgylyfpyklvkaadannvgllq 326
 |||||
 QY 302 YKLNGNVQOQVEFATSTANNTTANPTPAVDEIKVAK 338
 |||||
 DB 327 yklngnvpqvefatstsannttanptpavdeikvak 363
 |||||
 RESULT 6
 AAR63227
 ID AAR63227 standard; Protein; 368 AA.
 XX AAR63227;
 AC AAR63227;
 DT 23-JUN-1995 (first entry)
 XX
 DE Mycoplasma gallisepticum 40kd antigen.
 XX
 KW recombinant avipox virus; live vaccine; mycoplasma 40kd antigen;
 KW TTM-1.
 XX
 OS Mycoplasma gallisepticum.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 262 /note= "corresponds to a NNN codon"
 FT Misc-difference 283 /note= "corresponds to a NNN codon"
 FT
 XX W09423019-A.
 PN
 XX
 PD 13-OCT-1994.
 XX
 PF 31-MAR-1994; 94WO-JP00541.
 XX
 PR 31-MAR-1993; 93JP-0074139.
 PR 30-SEP-1993; 93JP-0245625.
 XX
 PA (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX
 PI Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;
 PI Saeki S, Saitos, Takahashi K;
 XX
 DR WPI; 1994-333181/41.
 DR N-PSDB; AAR63227.
 XX
 PT Recombinant avipox virus combining DNA encoding a polypeptide -
 PT exhibiting antigenicity of mycoplasma, useful for the production
 PT of a live vaccine
 XX
 PS Claim 4; Page 71-74; 123pp; Japanese.
 XX
 CC The plasmid pUTM-1P contains a sequence (the TTM-1 gene) coding
 CC for the 40kd antigen of Mycoplasma gallisepticum under the control
 CC of a synthetic promoter. A 1300 bp restriction fragment containing
 CC the promoter-ORF sequence was excised and was used in the
 CC construction of plasmid pNZ7929-R2. This in turn was involved in the
 CC construction of a recombinant avipox virus vector comprising the
 CC TTM-1 gene, DNA encoding the signal membrane anchor peptide from
 CC Newcastle Disease Virus haemagglutinin neuraminidase and FpV
 CC sequences. The recombinant avipox virus is useful as a live vaccine

CC to protect against infection by Mycoplasma gallisepticum.

XX
SQ Sequence 368 AA;

Query Match 80.4%; Score 1612; DB 15; Length 368;
Best Local Similarity 95.0%; Pred. No. 4.8e-102;
Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSEATVN 61
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 27 cmsitkdkanpnngqtqlearmeltdlinakartlasldgyakleaslsayseatvn 86
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 62 NNLNATLQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 87 mnlnatleqlkmaktnlesainqantdkttfdnehpnlveaykalkttleqratnlela 146
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 122 STAYNQIRNNLVLYNNKASSLITKTLDPNGTLLDSNEITANKNINNTLSTINEQKTN 181
-Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 147 staynqirnnlvlynnassliktldplnggmllidsneittvnrnintlstineqktn 206
QY 182 ADALNSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTYWNGDE 241
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 207 adalnsfkkviqneqsfvgftfnanvqpsnysfvafsadvtvnykyarrtywngde 266
QY 242 PSSRLANTNSTIDVSWIYSLAGTNTKYQFSNTPGPGTGYLYFFPKLVKAADANNVGLQ 301
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 267 psrllantnsitdvsxiyslagtnkyqfsnypgstgylyfpyklvkaadannvlgq 326
QY 302 YKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAK 338
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 327 yklnngnvqvefatstsannttanptqqlmrklilk 363

RESULT 7

AAAR76955
ID AAR76955 standard; Protein; 368 AA.
XX
AC AAR76955;
XX
DT 12-MAR-1996 (first entry)
XX
DE Mycoplasma gallisepticum antigenic protein TTM-1.
XX
KW Antigenic protein; vaccine; poultry; diagnosis; TTM-1.
XX
OS Mycoplasma gallisepticum.
XX
FH Key Location/Qualifiers
FT Misc-difference 262
FT /note= "any amino acid"
FT Misc-difference 283
FT /note= "any amino acid"

XX JP07133295-A.

XX 23-MAY-1995.

XX 27-AUG-1993; 93JP-0213102.

XX 27-AUG-1993; 93JP-0213102.

XX (SHIO) SHIONOGI & CO LTD.

XX WPI; 1995-220782/29.

DR N-PSDB; AAQ94711.

XX A new antigenic protein which reacts with Mycoplasma gallisepticum -
PT is useful in a component vaccine for use against poultry infected
PT with M. gallisepticum.

XX Claim 6; Figs 5-6; 33pp; Japanese.

XX

CC AAQ94711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein
CC TTM-1. TTM-1 can be used as a vaccine for M. gallisepticum
CC infectious diseases in poultry, and as a diagnostic agent for
CC M. gallisepticum.

XX
SQ Sequence 368 AA;

Query Match 80.0%; Score 1604; DB 16; Length 368;
Best Local Similarity 94.4%; Pred. No. 1.7e-101;
Matches 318; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKTEASLSAYSEATVN 61
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 27 cmsitkdkanpnngqtqlearmeltdlinakartlasldgyakleaslsayseatvn 86
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 62 NNLNATLQKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 87 mnlnatleqlkmaktnlesainqantdkttfdnehpnlvqpykalkttleqratnlela 146
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 122 STAYNQIRNNLVLYNNKASSLITKTLDPNGTLLDSNEITANKNINNTLSTINEQKTN 181
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 147 staynqirnnlvlynnassliktldplnggmllidsneittvnrnintlstineqktn 206
QY 182 ADALNSFKKVIQNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTYWNGDE 241
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 207 adalnsfkkviqneqsfvgftfnanvqpsnysfvafsadvtvnykyarrtywngde 266
QY 242 PSSRLANTNSTIDVSWIYSLAGTNTKYQFSNTPGPGTGYLYFFPKLVKAADANNVGLQ 301
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 267 psrllantnsitdvsxiyslagtnkyqfsnypgstgylyfpyklvkaadannvlgq 326
QY 302 YKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAK 338
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 327 yklnngnvqvefatstsannttanptqqlmrklilk 363

RESULT 8

AAAP93646
ID AAR93646 standard; protein; 235 AA.
XX

AC AAP93646;

XX 11-MAY-1990 (first entry)

XX Amino acid sequence of Mycoplasma gallisepticum (MGI) polypeptide.

XX Mycoplasma gallisepticum; Poultry vaccine; ss;

XX Mycoplasma gallisepticum.

XX EP345021-A.

XX 06-DEC-1989.

XX 31-MAY-1989; 89EP-0005441.

XX 02-JUN-1988; 88JP-0136343.

XX (JAPG) NIPPON ZEON KK.

PA (SHIO) SHIONOGI SEIVAKU KABUSHIKI KAISHA.

PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;

XX WPI; 1989-358393/49.

DR N-PSDB; AAN92568.

XX New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PT vaccines.

XX Disclosure; Fig.1a; 3lpp; English.

XX This amino acid sequence of MGI is encoded by M1 DNA and elicits an

CC antigen-antibody reaction with anti-MG poultry sera. It can be used as a
 CC vaccine to prevent and diagnose MG infection. Doseage is at least 1 micro
 CC gram vaccine / kg. No acute toxicity was noted with a dose of 5 mg / kg.
 XX
 SQ Sequence 235 AA;

Query Match 57.0%; Score 1142; DB 10; Length 235;
 Best Local Similarity 97.0%; Pred. No. 2.8e-70;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGOTQLEAARMELTDLINAKAMTILASLDYAKIEASISSAYSEATVN 61
 DB 1 cmsitkdkanpnnggtqlqaarmeltdlinakartilasldyakiessaysaeatvn 60
 QY 62 NNINATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLS 121
 DB 61 nninatleqlkmaaktulesainqantdkttfdnehpnlveaykalkttleqratnlegla 120
 QY 122 STAYNQIRNNLDVLYNKASSLITKTLDPNGGTLPLDLSNEITTTANKNNINNTLSTINEQKTN 181
 DB 121 staynqirnnldvlynassliktdlplnggmlldgsneittvnrninntlstineqktn 180
 QY 182 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSVAFSADVTPVNYKYARETV 236
 DB 181 adalsnfsikkvionneqsfvgtftnanvqpsnysvafsadvtpvnykyarrtv 235

RESULT 9
 AAR05081
 ID AAR05081 standard; protein; 235 AA.
 AC AAR05081;
 XX
 DT 08-OCT-1990 (first entry)
 XX
 DE MG-1 antigen.
 XX
 KW Mycoplasma gallisepticum; poultry; vaccine.
 XX
 PN JP02111795-A.
 XX
 PD 24-APR-1990.
 XX
 PF 02-JUN-1989; 89JP-0136343.
 XX
 PR 02-JUN-1989; 89JP-0136343.
 XX
 PA (JAPG) NIPPON ZEON KK.
 XX
 PA (SHIO) SHIONOGI KK.
 XX
 DR WPI; 1990-169109/22.
 DR N-PSDB; AAQ04686.
 XX
 PT Diagnostic and vaccine for poultry mycoplasma serum - utilises
 PT antigen protein of the disease and recombinant vector
 PT incorporated with its coding gene.
 XX
 PS Claim 2; Fig 1a; 20pp; Japanese.
 XX

CC DNA encoding the protein can be inserted into an expression vector
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also
 CC be ligated to other DNA to produce fusion proteins with an N-terminal
 CC bacterial enzyme sequence.
 CC See also AAR05081-2 and AAR06437-41.
 XX
 SQ Sequence 235 AA;

Query Match 57.0%; Score 1142; DB 11; Length 235;
 Best Local Similarity 97.0%; Pred. No. 2.8e-70;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGOTQLEAARMELTDLINAKAMTILASLDYAKIEASISSAYSEATVN 61
 DB 1 cmsitkdkanpnnggtqlqaarmeltdlinakartilasldyakiessaysaeatvn 60
 QY 62 NNINATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLS 121
 DB 61 nninatleqlkmaaktulesainqantdkttfdnehpnlveaykalkttleqratnlegla 120
 QY 122 STAYNQIRNNLDVLYNKASSLITKTLDPNGGTLPLDLSNEITTTANKNNINNTLSTINEQKTN 181
 DB 121 staynqirnnldvlynassliktdlplnggmlldgsneittvnrninntlstineqktn 180
 QY 182 ADALNSFIKKVIONNEQSFVGTFTNANVQPSNYSVAFSADVTPVNYKYARETV 236
 DB 181 adalsnfsikkvionneqsfvgtftnanvqpsnysvafsadvtpvnykyarrtv 235

RESULT 10
 AAR05082
 ID AAR05082 standard; protein; 261 AA.
 XX
 AC AAR05082;
 XX
 DT 08-OCT-1990 (first entry)
 XX
 DE TMG-1 antigen.
 XX
 KW Mycoplasma gallisepticum; poultry; vaccine.
 XX
 PN JP02111795-A.
 XX
 PD 24-APR-1990.
 XX
 PF 02-JUN-1989; 89JP-0136343.
 XX
 PR 02-JUN-1989; 89JP-0136343.
 XX
 PA (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI KK.
 XX
 DR WPI; 1990-169109/22.
 DR N-PSDB; AAQ04687.
 XX
 PT Diagnostic and vaccine for poultry mycoplasma serum - utilises
 PT antigen protein of the disease and recombinant vector
 PT incorporated with its coding gene.
 XX
 PS Claim 2; Fig 2; 20pp; Japanese.
 XX

CC DNA encoding the protein can be inserted into an expression vector
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also
 CC be ligated to other DNA to produce fusion proteins with an N-terminal
 CC bacterial enzyme sequence.
 CC See also AAR05081 and AAR06437-41.
 XX
 SQ Sequence 261 AA;

Query Match 57.0%; Score 1142; DB 11; Length 261;
 Best Local Similarity 97.0%; Pred. No. 3.2e-70;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 CMSITKDKANPNNGOTQLEAARMELTDLINAKAMTILASLDYAKIEASISSAYSEATVN 61
 DB 27 cmsitkdkanpnnggtqlqaarmeltdlinakartilasldyakiessaysaeatvn 86
 QY 62 NNINATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLS 121
 DB 87 nninatleqlkmaaktulesainqantdkttfdnehpnlveaykalkttleqratnlegla 146
 QY 122 STAYNQIRNNLDVLYNKASSLITKTLDPNGGTLPLDLSNEITTTANKNNINNTLSTINEQKTN 181

Db 147 staynqirnlvldynnasslittkldplngmildsneittvnrnntlstineqktn 206
 QY 182 ADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 236
 Db 207 adalsnsfikkvionneqsfvgtftnanvqpsnysfvafsadvtvnykyarrtv 261

RESULT 11

AA79911
 ID AAR79911 standard; Protein; 261 AA.
 AC AAR79911;
 XX
 XX 19-JUL-1996 (first entry)
 DE M.gallisepticum 261 amino acid protein.
 XX
 KW Detection; probe; primer; PCR; amplification; secretion; lung;
 KW avian chronic respiratory disease; respiratory tract; nasal cavity.
 XX
 OS Mycoplasma gallisepticum.
 XX
 PN JP07236498-A.
 XX
 PD 12-SEP-1995.
 XX
 PF 25-FEB-1994; 94JP-0052764.
 XX
 PR 25-FEB-1994; 94JP-0052764.
 XX
 PA (JAPG) NIPPON ZEON KK.
 PA (SHIO) SHIONOGI & CO LTD.
 XX
 DR WPI; 1995-347462/45.
 DR N-PSDB; AAT04076.
 XX

XX Detection of Mycoplasma gallisepticum - for the quick detection,
 PT i.e. within one day, of avian chronic respiratory diseases
 XX

PS Claim 3; Page 10-11; 1lpp; Japanese.

XX This is the amino acid sequence of a 261 amino acid protein encoded
 CC by a fragment of the Mycoplasma gallisepticum genome. The encoding
 CC sequence and the sequence of AAT04075 (encoding a 661 amino acid
 CC protein) can be used to detect M.gallisepticum using probes based on
 CC nucleotides 1125-1648 and primers based on nucleotides 449-466, the
 CC complement of bases 893-919, 1908-1934 and the complement of bases
 CC 2184-2210 of AAT04075 and a probe based on nucleotides 718-41 of the DNA
 CC encoding this protein. The method using these sequences is faster i.e. is
 CC able to detect M.gallisepticum, which causes avian chronic respiratory
 CC diseases, within one day, from avian secretions, washings from the lung,
 CC respiratory tract, nasal cavity, etc.
 XX

SQ Sequence 261 AA;

Query Match 57.08; Score 1142; DB 16; Length 261;
 Best Local Similarity 97.08; Pred. No. 3.2e-70;
 Matches 228; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 CMSITTKDANPNNGQTOLAEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 61
 Db 27 cmsitkddanpnngqtlqaarmeltdlinakartlasldgyakieslsayseaeavn 86

QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
 Db 87 nlnnatleqlkmaktnlesainqantdkttfdnehpnlveaykalkttleqratnleqla 146

QY 122 STAYNQIRNLVLDYNNKASSLITTKTLDPLNGTLLDSNEITTANKNINNTLSTINEOKTN 181
 Db 147 staynqirnlvldynnasslittkldplngmildsneittvnrnntlstineqktn 206

QY 182 ADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 236
 Db 207 adalsnsfikkvionneqsfvgtftnanvqpsnysfvafsadvtvnykyarrtv 261

RESULT 12

AA93959
 ID AAP93959 standard; protein; 261 AA.
 XX
 AC AAP93959;
 XX
 XX 11-MAY-1990 (first entry)
 DE Amino acid (AA) sequence of TMG-1 polypeptide.
 XX
 KW Mycoplasma gallisepticum; Poultry vaccine; ss;
 XX
 OS Mycoplasma gallisepticum.
 XX
 PN EP345021-A.
 XX
 PD 06-DEC-1989.
 XX
 PF 31-MAY-1989; 89EP-0005441.
 XX
 PR 02-JUN-1988; 88JP-0136343.
 XX

PA (JAPG) NIPPON ZEON KK; (SHIO) SHIONOGI SEIYAKU KABUSHIKI KAISHA.

XX Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;

XX WPI; 1989-358393/49.

XX N-PSDB; AAN92574.

XX New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
 PT vaccines.
 XX

PS Disclosure; Fig.2; 3lpp; English.

XX This AA sequence of TMG-1 is encoded by TM-1 base sequence. It has the
 CC same sequence as that of a polypeptide expressed in Mycoplasma
 CC gallisepticum in nature. When the corresponding DNA sequence is inserted
 CC into a recombinant vector used to transform a host the antigen protein
 CC produced can be used as a vaccine to prevent and diagnose MG infection.

XX SQ Sequence 261 AA;

Query Match 55.78; Score 1117; DB 10; Length 261;
 Best Local Similarity 96.28; Pred. No. 1.6e-68;
 Matches 227; Conservative 3; Mismatches 4; Indels 2; Gaps 2;

QY 2 CMSITTKDANPNNGQTOLAEARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 61
 Db 27 cmsitkddanpnngqtlqaarmeltdlinakartlasldgyakieslsayseaeavn 86

QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
 Db 87 nlnnatleqlkmaktnlesainqantdkttfdnehpnlveaykalkttleqratnleqla 146

QY 122 STAYNQIRNLVLDYNNKASSLITTKTLDPLNGTLLDSNEITTANKNINNTLSTINE-OKT 180
 Db 147 staynqirnlvldynnasslittkldplngmildsneittvnrn-ntlstineqktn 205

QY 181 NADALSNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARRTV 236
 Db 206 nadalsnsfikkvionneqsfvgtftnanvqpsnysfvafsadvtvnykyarrtv 261

RESULT 13

AA63226
 ID AAR63226 standard; Protein; 661 AA.
 XX

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:49:08 ; Search time 22.56 seconds

(without alignments)
426.582 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086

Perfect score: 2004

Sequence: 1 GCMSTTKKDANPNNGTQLE.....SSNNADKIPGRRPGTEL 394

-Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

*Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCRTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1910	95.3	615	2	US-08-525-742-10
2	1856	92.6	610	2	US-08-525-742-8
3	1638	81.7	368	2	US-08-525-742-4
4	1612	80.4	368	1	US-08-185-851A-4
5	806	40.2	661	2	US-08-525-742-2
6	159	7.9	933	3	US-08-293-728-2
7	159	7.9	933	3	US-08-421-868-2
8	136	6.8	1098	4	US-08-923-992A-8
9	135.5	6.8	1002	4	US-09-268-347-24
10	135.5	6.8	1073	4	US-09-541-782-6
11	134	6.7	1004	4	US-09-268-347-30
12	131.5	6.6	1164	4	US-08-923-992A-2
13	131	6.5	1104	4	US-08-923-992A-4
14	129	6.4	2048	4	US-09-268-347-48
15	126.5	6.3	1128	4	US-08-923-992A-6
16	126.5	6.3	1164	4	US-08-923-992A-10
17	125.5	6.3	518	3	US-09-043-123-2
18	125.5	6.3	3788	4	US-09-336-447A-76
19	124.5	6.2	2411	4	US-09-268-347-36
20	122.5	6.1	1536	1	US-09-038-682-2
21	122.5	6.1	1536	1	US-08-302-832-2
22	122.5	6.1	1536	2	US-08-530-198-2
23	122.5	6.1	1536	2	US-08-469-880-2
24	122.5	6.1	1536	2	US-08-728-470-2
25	122.5	6.1	1536	2	US-08-617-697-2
26	122.5	6.1	1536	4	US-08-719-641-2
27	122.5	6.1	1565	6	5352450-2
28					Sequence 10, Appl
29					Sequence 8, Appl
30					Sequence 4, Appl
31					Sequence 2, Appl
32					Sequence 2, Appl
33					Sequence 2, Appl
34					Sequence 2, Appl
35					Sequence 2, Appl
36					Sequence 2, Appl
37					Sequence 2, Appl
38					Sequence 2, Appl
39					Sequence 2, Appl
40					Sequence 2, Appl
41					Sequence 2, Appl
42					Sequence 2, Appl
43					Sequence 2, Appl
44					Sequence 2, Appl
45					Sequence 2, Appl

Sequence 2, Appl
Sequence 28, Appl
Sequence 34, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 49, Appl
Sequence 47, Appl
Sequence 4, Appl
Sequence 2, Appl
Sequence 15, Appl
Sequence 4, Appl
Sequence 33, Appl
Sequence 4, Appl
Sequence 33, Appl
Sequence 33, Appl
Sequence 9, Appl
Sequence 9, Appl

ALIGNMENTS

RESULT 1

US-08-525-742-10

; Sequence 10, Application US/08525742

; Patent No. 5871742

; GENERAL INFORMATION:

; APPLICANT: Saito, Shuji

; APPLICANT: Ohkawa, Setsuko

; APPLICANT: Saeki, Sakiko

; APPLICANT: Ohsawa, Ikuroh

; APPLICANT: Funato, Hirono

; APPLICANT: Iritani, Yoshikazu

; APPLICANT: Aoyama, Shigeml

; APPLICANT: Takahashi, Kiyoohito

; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE

; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND

; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &

; STREET: 1725 K Street, Suite 1000

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20006

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/525,742

; FILING DATE: 25-SEP-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 05-074139

; FILING DATE: 31-MAR-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 05-245625

; FILING DATE: 30-SEP-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/JP94/00541

; FILING DATE: 31-MAR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Mclelland, Le-Nhung

; REGISTRATION NUMBER: 31,541

; REFERENCE/DOCKET NUMBER: 950811

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-659-2930

TELEFAX: 202-8870357
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 615 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-525-742-10

Query Match 95.3%; Score 1910; DB 2; Length 615;
Best Local Similarity 98.2%; Pred. No. 3.9e-132; Indels 0; Gaps 0;
Matches 376; Conservative 3; Mismatches 4;

QY 2 CMSITKDDANPNNGOTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSAETVN 61
DB 27 CMSITKDDANPNNGOTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSAETVN 86
QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLYEAYKALKTTLEQRATNLEGLS 121
DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLYEAYKALKTTLEQRATNLEGLA 146
QY 122 STAYNOIRNVLVDLYNKASLLTKTLDPLNGGTLDSNEITTTANKNINNTLSTINEOKTN 181
DB 147 STAYNOIRNVLVDLYNKASLLTKTLDPLNGGTLDSNEITTTANKNINNTLSTINEOKTN 206
QY 182 ADALSNSEFKKVIQNNQESFVGTFTNANQVPSNYSFVAFSADVTPVNYKYARRTVWNGDE 241
DB 207 ADALSNSEFKKVIQNNQESFVGTFTNANQVPSNYSFVAFSADVTPVNYKYARRTVWNGDE 266
QY 242 PSSRILANTNSITDVSNIWYSLAGTNTKYQFSNYPSTGYLFFPKLYKAADANNVGLQ 301
DB 267 PSSRILANTNSITDVSNIWYSLAGTNTKYQFSNYPSTGYLFFPKLYKAADANNVGLQ 326
QY 302 YKLNGNVQVVEFATSTANNTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 361
DB 327 YKLNGNVQVVEFATSTANNTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 386
QY 362 MNKVAPMIGNIYLSNENNAKDI 384
DB 387 MNKVAPMIGNIYLSNENNAKDI 409

RESULT 2
US-08-525-742-8
Sequence 8, Application US/08525742
Patent No. 5871742
GENERAL INFORMATION:
APPLICANT: Saito, Shuji
APPLICANT: Ohkawa, Setsuko
APPLICANT: Saeki, Sakiko
APPLICANT: Ohsawa, Ikuroh
APPLICANT: Funato, Hirono
APPLICANT: Iritani, Yoshikazu
APPLICANT: Aoyama, Shigemi
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
TITLE OF INVENTION: AS USE THEREOF
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
ADDRESSEE: NAUGHTON
STREET: 1725 K Street, Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,742
FILING DATE: 25-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-074139
FILING DATE: 31-MAR-1993
APPLICATION DATA: JP 05-245625
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/00541
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Mclelland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 950811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-8870357
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-525-742-8

Query Match 92.6%; Score 1856; DB 2; Length 610;
Best Local Similarity 95.5%; Pred. No. 3.5e-128; Indels 0; Gaps 0;
Matches 365; Conservative 10; Mismatches 7;

QY 2 CMSITKDDANPNNGOTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSAETVN 61
DB 27 CMSITKDDANPNNGOTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSAETVN 86
QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLYEAYKALKTTLEQRATNLEGLS 121
DB 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLYEAYKALKTTLEQRATNLEGLS 146
QY 122 STAYNOIRNVLVDLYNKASLLTKTLDPLNGGTLDSNEITTTANKNINNTLSTINEOKTN 181
DB 147 STAYNOIRNVLVDLYNKASLLTKTLDPLNGGTLDSNEITTTANKNINNTLSTINEOKTN 206
QY 182 ADALSNSEFKKVIQNNQESFVGTFTNANQVPSNYSFVAFSADVTPVNYKYARRTVWNGDE 241
DB 207 ADALSNSEFKKVIQNNQESFVGTFTNANQVPSNYSFVAFSADVTPVNYKYARRTVWNGDE 266
QY 242 PSSRILANTNSITDVSNIWYSLAGTNTKYQFSNYPSTGYLFFPKLYKAADANNVGLQ 301
DB 267 PSSRILANTNSITDVSNIWYSLAGTNTKYQFSNYPSTGYLFFPKLYKAADANNVGLQ 326
QY 302 YKLNGNVQVVEFATSTANNTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 361
DB 327 YKLNGNVQVVEFATSTANNTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGN 386
QY 362 MNKVAPMIGNIYLSNENNAKDI 383
DB 387 MNKVAPMIGNIYLSNENNAKDI 408

RESULT 3
US-08-525-742-4
Sequence 4, Application US/08525742
Patent No. 5871742
GENERAL INFORMATION:
APPLICANT: Saito, Shuji
APPLICANT: Ohkawa, Setsuko
APPLICANT: Saeki, Sakiko
APPLICANT: Ohsawa, Ikuroh
APPLICANT: Funato, Hirono

APPLICANT: Iritani, Yoshikazu
APPLICANT: Aoyama, Shigemi
APPLICANT: Takahashi, Kiyochoito
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
TITLE OF INVENTION: AS USE THEREOF
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
ADDRESSEE: NAUGHTON
STREET: 1725 K Street, Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/525.742
FILING DATE: 25-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-074139
FILING DATE: 31-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-245625
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/00541
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: McLeLland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 950811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-8870357
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-525-742-4

Query Match 81.7%; Score 1638; DB 2; Length 368;
Best Local Similarity 95.5%; Pred. No. 1.6e-112; Indels 0; Gaps 0;
Matches 322; Conservative -6; Mismatches 9;
Qy 2 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 61
Db 27 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 86
Qy 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
Qy 122 STAYNQIRNNLVLDYKNASSLITKTLDPNGTGLDSDNEITTNANKNINNTLSTINEQKTN 181
Db 147 STAYNQIRNNLVLDYKNASSLITKTLDPNGTGLDSDNEITTNANKNINNTLSTINEQKTN 206
Qy 182 ADALNSFIKKVQIONNEQSFVGTFTNANQPSNYSFVAFSDVTPVNYKYARRTVWNGDE 241
Db 207 ADALNSFIKKVQIONNEQSFVGTFTNANQPSNYSFVAFSDVTPVNYKYARRTVWNGDE 266
Qy 242 PSSRLANTNSTIDVSWIYSYLAGTNTKYQSFNSYGPSTGYLYFPYKLVKADANNVGLQ 301
Db 267 PSSRLANTNSTIDVSWIYSYLAGTNTKYQSFNSYGPSTGYLYFPYKLVKADANNVGLQ 326

Qy 302 YKLNGNQQVEFATSTSANNTANTPTPAVDKIVAK 338
Db 327 YKLNGNQQVEFATSTSANNTANTPTQOLMLKLLK 363

RESULT 4

US-08-185-851A-4
Sequence 4, Application US/08185851A
Patent No. 5489430
GENERAL INFORMATION:
APPLICANT: Saito, Shuji
APPLICANT: Ohkawa, Setsuko
APPLICANT: Fujisawa, Ayumi
APPLICANT: Iritani, Yoshikazu
APPLICANT: Aoyama, Shigemi
TITLE OF INVENTION: Poultry Mycoplasma Antigens, Gene
TITLE OF INVENTION: Thereof and Recombinant Vectors Containing the Gene As
TITLE OF INVENTION: Well As Vaccines Utilizing the Same
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, McLeLland &
ADDRESSEE: Naughton
STREET: 1725 K Street, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS version 5.0
SOFTWARE: ASCII from Word Perfect version 5.1
CURRENT APPLICATION DATA: US/08/185.851A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Theresa M. Stevens-Smith
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: PO-8-A930918
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-887-0357
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-185-851A-4

Query Match 80.4%; Score 1612; DB 1; Length 368;
Best Local Similarity 95.0%; Pred. No. 1.3e-110;
Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
Qy 2 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 61
Db 27 CMSITKDKANPNNGQTQLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN 86
Qy 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db 87 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146
Qy 122 STAYNQIRNNLVLDYKNASSLITKTLDPNGTGLDSDNEITTNANKNINNTLSTINEQKTN 181
Db 147 STAYNQIRNNLVLDYKNASSLITKTLDPNGTGLDSDNEITTNANKNINNTLSTINEQKTN 206
Qy 182 ADALNSFIKKVQIONNEQSFVGTFTNANQPSNYSFVAFSDVTPVNYKYARRTVWNGDE 241
Db 207 ADALNSFIKKVQIONNEQSFVGTFTNANQPSNYSFVAFSDVTPVNYKYARRTVWNGDE 266

242 PSSRLANTNSITDVSIIYSLAGTNTKYQSFSGNYGPGSTGYLYPPYKLVKAAADANNVGLQ 301
 267 PSSRLANTNSITDVSIIYSLAGTNTKYQSFSGNYGPGSTGYLYPPYKLVKAAADANNVGLQ 326
 302 YKLNNGNVQVEFEATSTANNTANPTPAVDKIVAK 338
 327 YKLNNGNVQVEFEATSTANNTANPTQQLMLKLLK 363

RESULT 5
 US-08-525-742-2
 ; Sequence 2, Application US/08525742
 ; Patent No. 5871742
 ; GENERAL INFORMATION:
 ; APPLICANT: Saito, Shuji
 ; APPLICANT: Ohkawa, Setsuko
 ; APPLICANT: Saeki, Sakiko
 ; APPLICANT: Ohsawa, Ikuroh
 ; APPLICANT: Funato, Hiroko
 ; APPLICANT: Iritani, Yoshikazu
 ; APPLICANT: Aoyama, Shigemi
 ; APPLICANT: Takahashi, Kiyochito
 ; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
 ; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
 ; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
 ; TITLE OF INVENTION: AS USE THEREOF
 ; NUMBER OF SEQUENCES: 51
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &
 ; STREET: 1725 K Street, Suite 1000
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/525,742
 ; FILING DATE: 25-SEP-1995
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 05-074139
 ; FILING DATE: 31-MAR-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 05-245625
 ; FILING DATE: 30-SEP-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/JP94/00541
 ; FILING DATE: 31-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mclelland, Le-Nhung
 ; REGISTRATION NUMBER: 31,541
 ; REFERENCE/DOCKET NUMBER: 950811
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-659-2930
 ; TELEFAX: 202-8870357
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 661 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-525-742-2

Query Match 40.2%; Score 806; DB 2; Length 661;
 Best Local Similarity 42.8%; Pred. No. 2.9e-51;
 Matches 185; Conservative 62; Mismatches 123; Indels 62; Gaps 12;

9 DANNPNCOTLEARMELTD-----LINAKAMTSLASLDYAKIEASLSAYSEATVN 61
 45 NTPNSDQGMNAAAKELADAKAALTTLINGETANLASVEDYAKIKSELTSAYETAKAVS 104
 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGLS 121
 105 AKTGATLNEVNEAKTTLDAAIKKAASAKNDFDAHGSLVEAYNNLKETLKEEKNLDSLA 164
 122 STAYNQIRNNLVLDLYNKASSLITKTLDPLAGGTLTLLDSNEITANKNNINNTLSTINEQKN 181
 165 NENYAAIRTNLSLYEKANTIVTATLDPAT-GNIPYMSVTOANQDITNATSRILIAKQN 223
 182 ADALNSFIKKYVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTTPVNYKYARTVWNGDE 241
 224 ADNLANSFIKQSLVKNNLITRDVANNQOE-QPANYSFVGFVSVNVDTPNWNFAQRKVASEN 282
 242 -----PSSRLANTNSITDVSIIYSLAGTNTKYQSFSGNYGPGSTGYLYPPYKLVKAAAD 294
 283 TPLATTPAEDATQQAASLTDSVSIYSLNGAEAKYTLSPFVGAETKATLYLYPPYKLVKTS 342
 295 ANNVLQYKLNNGNVQVEF-----ATSTSAN-NTTANP-----TPAYDEIKVAKIV 340
 343 --NVGLQYKLNGBGDTKQINFVOTPASGSSDVAANEETMASPAEQSNAPTVDYDDIKIAKVA 400
 341 LSGLRFGONTIELSVPTGEGNNKVPAPMIGNIYLS-----NENN 380
 401 LSNLKFNSNTIEFSVPTG-----KAAPMIGNMILTSSSEVKNKNIYDGLFGSNFNENN 455
 381 -----ADKIPGY 387
 456 PTAVTVDLLKGY 467

RESULT 6
 US-08-293-728-2
 ; Sequence 2, Application US/08293728D
 ; Patent No. 6008341
 ; GENERAL INFORMATION:
 ; APPLICANT: Foster, Timothy J.
 ; APPLICANT: McDevitt, Damien L.
 ; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
 ; FILE REFERENCE: 05344.105011
 ; CURRENT APPLICATION NUMBER: US/08/293,728D
 ; CURRENT FILING DATE: 1994-08-22
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 933
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 ; US-08-293-728-2

Query Match 7.9%; Score 159; DB 3; Length 933;
 Best Local Similarity 22.8%; Pred. No. 0.0011;
 Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;

3 MSTTKDANPNNGOTOL--EARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATV 60
 73 VSDTKSSNTNNGETSVQNPAAQETQSSSTNATT-----EETPTVGEATT 120
 61 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRTNLEGL 120
 121 TTQANTPATQSSNTNAEELVNO-TSNETTFND--TNTVSSVNS-----PQNSTNAENV 172
 121 SST-----AYNQIRNNLVLDLYNK-----ASSLITKTLDPLNGTLL 156
 173 STTQDTSTEATPSNNEASPOSTDASNKDVVNQAVNTSAPMRAFSLAAVAADAPAACTDI 232
 157 DSNEITTANKNNINNTLSTINEQ-----KTN-ADALNSNFIK-----KVTONNEQSFVGTFTN 207
 233 -TNQLTNTVTCIDSGTIVYHQAGYVVKLVNGYFVSPNSAVKGDGTFKITVPKRLNNGVTST 291

Qy 208 ANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRILANTNSITDVSIIYSLAG-TN 266
Db 292 AKVPP-----IMAGDQ-----VLANGVIDSGNVIYFTDYN 324
Qy 267 TKYQFSNYSFGPGYLYFPKLVKAADANNVQYKLNNGNVQVQVEFATSTANNITAN 326
Db 325 TKDDVKATLTMPA--YI-----DPENV-----KKTGNV-----TLATGIGSTTAN 362
Qy 327 PTPAVDEIKVAKI-----VLSGLRFQNTIELSVPTGEGNMKNKVPAMI-GNIYLSNE 378
Db 363 KTVLVDEYKGYKFFYNLSIKGTIDQIDKTNNTYRQTIYVNPSSGDNVIAPLVGLNKPNTDS 422
Qy 379 N 379
Db 423 N 423
RESULT 7
US-09-421-868-2
; Sequence 2, Application US/09421868
; Patent No. 6177084
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: McDevitt, Damien L.
; TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene
; FILE REFERENCE: 05344.105011
; CURRENT APPLICATION NUMBER: US/09/421.868
; CURRENT FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 08/293,728
; PRIOR FILING DATE: 1994-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 933
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-421-868-2

Query Match 7.9%; Score 159; DB 4; Length 933;
Best Local Similarity 22.8%; Pred. No. 0.0011;
Matches 96; Conservative 57; Mismatches 154; Indels 114; Gaps 20;
Qy 3 MSITKKDANPNNGTQL--EAARMELTDLINAKAMTLASLDYAKIEASLSAYSEACTV 60
Db 73 VSDTKTSNTNGETSVQAQNAQOETTQSSSTNATT-----EETPVTGEATTT 120
Qy 61 NNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGL 120
Db 121 TTNQANTPATYQSSNTAEELVQ--TSNETTFND--TNTVSSVNS-----PQNSTNAENV 172
Qy 121 SST-----AYNQIRNVLVDLYNK-----ASSLTKTLDPLNGTLL 156
Db 173 STTQDTSTPATPSNNESAQSTDSAKNDVQNAVNTSAPRMFASLAADAAPAAQTDI 232
Qy 157 DSNETTANKNNLTSTINEQ----KTN-ADALNSFIK----KVIONNQSFGVGTFTN 207
Db 233 -TNQLTNVTVGIDSGTIVYHQAGVKLVNYSFVPSNAVKGDTFKITVPKELNNGVTST 291
Qy 208 ANVQPSNYSFVAFSADVTPVNYKYARRTVWNGDEPSSRILANTNSITDVSIIYSLAG-TN 266
Db 292 AKVPP-----IMAGDQ-----VLANGVIDSGNVIYFTDYN 324
Qy 267 TKYQFSNYSFGPGYLYFPKLVKAADANNVQYKLNNGNVQVQVEFATSTANNITAN 326
Db 325 TKDDVKATLTMPA--YI-----DPENV-----KKTGNV-----TLATGIGSTTAN 362
Qy 327 PTPAVDEIKVAKI-----VLSGLRFQNTIELSVPTGEGNMKNKVPAMI-GNIYLSNE 378
Db 363 KTVLVDEYKGYKFFYNLSIKGTIDQIDKTNNTYRQTIYVNPSSGDNVIAPLVGLNKPNTDS 422
Qy 379 N 379

Db 423 N 423
RESULT 8
US-08-923-992A-8
; Sequence 8, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: NO. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923.992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-8

Query Match 6.8%; Score 136; DB 4; Length 1098;
Best Local Similarity 22.7%; Pred. No. 0.064;
Matches 84; Conservative 51; Mismatches 159; Indels 76; Gaps 14;
Qy 5 ITKDDANPNNGTQLEAARMELTDLINA-KAMTLASLDYAKIEASLSAYSE-----AE 58
Db 192 IRKQAQDPKKEDAEVKVREELKGLFSSTKAGLDQIEQHVKKTSSEENTQKVDEHYAN 251
Qy 59 TVNNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNL-----VEAYKALK 107
Db 252 SLQNLQAQKSELELDKATTN-----EQATQVKNOQFLENAQKLEIQLIKETNVKLYKAMS 306
Qy 108 TLEQRATNLEGLSSTAYNQIRNNLVNKNASSLITTTLDPLNGTLTLDNSNEITTANKN 167
Db 307 ESLEQVEKELKHNSA-----NLEDLVAKSKEIVREVEGKLNOSKNLP--ELQKLEEE 357
Qy 168 INNTLSTINEQ-----KTNADALNSFIKKVI---ONNEQSEVGTFTNANVQPSNYSEVA 219
Db 358 AHSKLVQVVEDFRKKFKTSEQVTPKKRLKRLDLAANNQOKI-----ELTVSPENI---- 408
Qy 220 PSADVTPVNYKYARTVWNGDEPSSRILANTNSITDVSIIYSLAGTNTKYQFSFS----- 274
Db 409 -----TVYEGEDVKFTVAKSDSKTTLDFSDLL-----TKYNPSVSDRIST 449
Qy 275 NYGPSTG---YLYPPYKLVKAADANNVQYKLNNGNVQVQVEFATSTANNITANP-TPA 330

Db 450 NYKTTNDNHKIAEITIKNLKNSQTVLTKAKDDSGNVVEKFTTITVQKKEKQVPTPE 509
 QY 331 VDEIKVAKIV 340
 Db 510 QKDSKTEKV 519

RESULT 9
 US-09-268-347-24
 ; Sequence 24, Application US/09268347
 ; Patent No. 6335182
 ; GENERAL INFORMATION:
 ; APPLICANT: Loosmore, Sheena M.
 ; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
 ; FILE REFERENCE: 1038-860
 ; CURRENT APPLICATION NUMBER: US/09/268,347
 ; CURRENT FILING DATE: 1999-03-16
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 24
 ; LENGTH: 1002
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-09-268-347-24

Query Match 6.8%; Score 135.5; DB 4; Length 1002;
 Best Local Similarity 19.2%; Pred. No. 0.061;
 Matches 94; Conservative 60; Mismatches 161; Indels 175; Gaps 21;

QY 1 GCMSTTKKDANPNNGOTQLEAARM-----ELTDLINAKAMTSLADYAKIEASLSAYSE 56
 Db 349 GKLAQKVLKVSANGINPYKISNVADGTEQDAVSEK--QLKALQD-KQVTLASNAYAN 405
 QY 57 -----AETVNNLNATLQKMAKTNLESA-----INQANTDKTTFDN 94
 Db 406 GGSADGAGKATQTLGDLNFKFKSTDSSELLNIKAAGDTVTFTPKGVSQVGDGKATIQD 465
 QY 95 -----EHPNLVEAYKAL-----KTTL----- 110
 Db 466 GAKTTGLVEASELVDLSNKLGHKVGKDGATGDTHTDITLVKSGDKVTLKAGDNLV 525
 QY 111 EQRATN-----LEGSLSTAYNQIRNNLVLYNKASSLIQK---TLDP---LNGGTL 156
 Db 526 KQETNFTYVLRDELTCVKSVKEDTENG-----NGASTKITKDGTLTITPANDANGAAAT 581
 QY 157 DSNEITANKNNINNTLSTINEQKTNADALSNSFIKKVIONNEOSFVGTETNANVQPSNYS 216
 Db 582 DADKIKVASDGI-----SAGNKAVKNV-----SGLKKFGDANFP----- 617
 QY 217 FVAFSADVTPVNYKYARRTVWNGDEPS-----SRILANTNSITDVSWIYS---LAG 264
 Db 618 -LTSADNLTKQYDNAYKGLTNLDEKSKGQTPVADNTAATVCDLRLGLGWISADKTTG 676
 QY 265 TNTKYQSFNSNYGPGSTGYLYFPYKLVKAADANNVGLQYKLNNGVQVVEFATSTSANNTT 324
 Db 677 ESKEYSAQVRNANE-----VFKSGNGINVSCKTLDNGTREITTEFLAKDEN--- 722
 QY 325 ANPTPAVDEIKVAKIVLSGLRFGQNTIEL---SVPTGEGNM-----NKVAPMI 369
 Db 723 -----ATAFGSGKALRDNRTVAIGTGNVNVNAEKSGAFGDPNIEDKA 764
 QY 370 GNYLSSNEN 379
 Db 765 GGSYAFGNDN 774

RESULT 10
 US-09-541-782-6
 ; Sequence 6, Application US/09541782
 ; Patent No. 6284480
 ; GENERAL INFORMATION:

; APPLICANT: Nislow, Corey
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Beraud, Christophe
 ; TITLE OF INVENTION: Antifungal Assay
 ; FILE REFERENCE: 1015
 ; CURRENT APPLICATION NUMBER: US/09/541,782
 ; CURRENT FILING DATE: 2000-04-03
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 1073
 ; TYPE: PRT
 ; ORGANISM: S.pombe
 US-09-541-782-6

Query Match 6.8%; Score 135.5; DB 4; Length 1073;
 Best Local Similarity 20.9%; Pred. No. 0.067;
 Matches 97; Conservative 76; Mismatches 182; Indels 109; Gaps 21;
 QY 14 NQO-TQLEAARMELTDLINAKA-MTLASL-----QDYAKIEASLSAYSEATVNNLNAT 67
 Db 611 NGYFTLLDNDFASMEELLNTHSNQLLISMKTITEHFQSLDEALQASCAVPSNLSLDLI 670
 QY 68 LEOLKMAKTNLESAINQANTDKTTFDNEHPN-----LVEAYKALKTT---LEORATNLEG 119
 Db 671 VSELADSKNSLLDALEHSLQDISMSQKLGISSELIELQDKMKESYRQLVQELRSLYN 730
 QY 120 LSTAYNQIRNNLVLYNKASSLIQKTLPLNGGTLDSNEIT-----TA 164
 Db 731 LQHTHEESQKELMYGVNRDIDALVKTCITSLNDADILSDQKSKFESKQOQLIANI 790
 QY 165 NKNINNTLSTINEQ-KTNADAL-----SNSFIKK---VIONNEQSFVGTETNANVQPSN 214
 Db 791 GKIVSNFLOEQNESLYTKADILHSLNDTNSNIRKANEIMNRSEEL---RNA----- 841
 QY 215 YSFVAFSADVTPVNYKYARRTVWNG-----DEPSSRILANTNSITDVSWIYSLA----- 263
 Db 842 -----ASQAEIVGANKERIOKTVENGSQLDSDSKSKAHSNRSMYD---HCLALAESQK 893
 QY 264 GNTKYQ-----FSSNYGPGSTGYLYFPYKLVKAADANNVGL---QYKLNNGVQ 310
 Db 894 GVNLEVTDLRLQKVKHESEDNTEKHOOL---LDLESVLGNNNDNLIDSITKPTHELQ 950
 QY 311 QVE-----FATSTSANNTT-----ANPTPAVDEIKVAKIVLSGLRFGQNTI----- 351
 Db 951 KITDHLVKGTTSLANTHNELLGDESCLNLETTIEDTSLVKLETTGDTPTSKRELPAFPS 1010
 QY 352 -----ELSVPTGEGNMKNVAPMIGNIYLSNENNADKIPGRRP 390
 Db 1011 WTRDSSLIKETTNLSDSKKVFREYITSSNQTNPDV---YDKP 1052

RESULT 11
 US-09-268-347-30
 ; Sequence 30, Application US/09268347
 ; Patent No. 6335182
 ; GENERAL INFORMATION:
 ; APPLICANT: Loosmore, Sheena M.
 ; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
 ; FILE REFERENCE: 1038-860
 ; CURRENT APPLICATION NUMBER: US/09/268,347
 ; CURRENT FILING DATE: 1999-03-16
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 30
 ; LENGTH: 1004
 ; TYPE: PRT
 ; ORGANISM: Haemophilus influenzae
 US-09-268-347-30

Query Match 6.7%; Score 134; DB 4; Length 1004;

Best Local Similarity 20.3%; Pred. No. 0.079;
Matches 100; Conservative 53; Mismatches 160; Indels 180; Gaps 22;

QY 1 GCMSTTKDANPNNGQTOLEARM---ELDTLNAKAMTASLDYAKIEASLSAYSE 56
Db 346 GKLAATKVLKVSANPNVKISNVADGENTDAVSFK--QKALOD-KOVLTSASNAYAN 402
QY 57 A-----ETVNNLNATLEQLKMAKTNLESAINQANTDKTTF-----DN-- 94
Db 403 GGSADGCKGIOTLSNGLN-----PKFKSTDGELLNKAENDTVTFTPKKGSVQVGGDK 457
QY 95 -----BHPNLVEAYKAL-----KTTL--- 110
Db 458 ATIQDGAKTTLGLVBASELVDLSNKLKWKVGTDGTGVTGTDHTDITLVKSGDKVTLKAG 517
QY 111 -----EQRATN-----LEGLSTAYNOIRNNLVLDLYNKASSLITK---TLDPNGST 154
Db 518 DNLKVKQGGTFTYALKDELIDVRSVERKDTANGA-----NGASTKITDKGLTITPANGAG 573
QY 155 LLDSEIITANKNINNTLSTINEQKTNADALSNSPIKKVIONNEQSFVGTFTNANVQPSN 214
Db 574 AAGA-----NTANTISVTKDGISA-GNKAVKNV-----SGLKKFGDANFDP-- 614
QY 215 YSFVAFSADVPVNVKYARRVWNGDEPS-----SRILANTNSITDVSWIYSLAGT 265
Db 615 ---LTSADNLTKQYDNAYKGLTNLDEKSGKQTPTVADNTAATVGLRGLGWISADKT 671
QY 266 ---NTKYQSFNSYGPSTGYLYFPYKLVKAADANNVGLQYKLNNGNVQVFEATSTSAN 321
Db 672 KGELNKEYNAQVRNANE-----VKFKSGNGINVSCKTLDNGTRETITFELAKDEN 720
QY 322 NTAAPTAVDEIKVAKIVLSGLRFGQNTIEL---SVPTGEGNM-----NKVA 366
Db 721 -----AIAFGSGSKALRDNVTAIGTGNVYNAEKSAGFGDPNIE 759
QY 367 PMIGNIYLSNEN 379
Db 760 DRAGGSYAFGNDN 772

RESULT 12
US-08-923-992A-2
; Sequence 2, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER:
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1164 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-2

Query Match 6.6%; Score 131.5; DB 4; Length 1164;
Best Local Similarity 22.8%; Pred. No. 0.15;
Matches 84; Conservative 51; Mismatches 157; Indels 77; Gaps 15;

QY 7 KKDANPNNGQTOLEA-ARMELTDLINA-KAMTLASLDYAKIEASLSAYSE-----AET 59
Db 230 RKQAOADKDEDAEVKVRRELKGLFSSTKAGLDQEIQEHVKKETSSEENTQKVDEHYANS 289
QY 60 VNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNL-----VEAYKALKT 108
Db 290 LONLAQKSLLELDKATTN-----EQATQVKNQFLENAQKLKEIQPLIKETNVKLYKAMSE 344
QY 109 TLEQRATNLEGLSSPAYNQIRNNLVLDLYNKASSLITKTLDPNGSTLLDSNEITTTANKNI 168
Db 345 SLEQVEKELKHNSA-----NLEDVAKSEIVREYEGKLNQKNLP--ELKQLEEEA 395
QY 169 NNTLSTINEQ-----KTNADALSNSFIKKVI---QNEQSFVGTFTNANVQPSNYSFVAF 220
Db 396 HSKLKQVVEDFRKFKTSEQVTPKRVKRDLAANENNOOKI-----ELTVSPENI 445
QY 221 SADVTPVNVKYARRVWNGDEPSRILANTNSITDVSWIYSLAGTNTKYQSF 275
Db 446 -----TWEGEDVKFTVAKSDSKTTLDFSDLL-----TKYNPSVSDRISIN 487
QY 276 YGPSTG---YLYFPYKLVKAADANNVGLQYKLNNGNVQVFEATSTSANNTTANP-TPAV 331
Db 488 YKTNNDNHKIAETITIKNLKLNESQVTVTLKAKDDSGNVVEKTTITVQKKEEQVKTPEQ 547
QY 332 DEIKVAKIV 340
Db 548 KDSKTEEKV 556

RESULT 13
US-08-923-992A-4
; Sequence 4, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; REFERENCE/DOCKET NUMBER: US 60/024,707

```

; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1104 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SS-08-923-992A-4

```

```

Query Match      6.5%; Score 131; DB 4; Length 1104;
Best Local Similarity 20.2%; Pred. No. 0.15;
Matches 85; Conservative 50; Mismatches 154; Indels 132; Gaps 17;

  QY      7  KKDANPNNGQTQLEA-ARMELTDLINA-KAMTLASLDQYAKIEASLSAYSE-----AET 59
  DB      199  RKQAAQAKKEDAEVKVREELGKLFSSTKAGLDQOIQEHVKKETSSEENTQKVDEHYANS 258

  QY      60  VNNNLNATLEQLMAKTNLESANQNTDFTFNEHPNL-----VEAYKALKT 108
  DB      259  LQNLQAQKSLBELDKAT'N-----EQATQVKNQFLFENAAQKLEKIOPLIKETNVKLYKAMSE 313

  QY      109  TLEQRATNLSEGLSTAYNQRNLLVDLYNKRASSLITKTLDPINGTGTLDSNEITTANKNI 168
  DB      314  SLEQVEKELAHNSEA-----NLQDLVAKSKEIVREVEGLNQSKNLP--ELAQLEEEA 364

  QY      169  NNTLSTINEQ-----KTNADALSFTKKVI---QNNESQVGVGTFTRANVOPSNYSFVAF 220
  DB      365  HSKLKQVVEHFRKFKFTSEQVTPKKRYKRDLAANNENQOKI-----ELTVSPENI-----414

  QY      221  SADVTPVNYKYARTVNVGDBPSSRIILANTNSITDWSWIYSLAGTNTKIQFS--FSNYGP 278
  DB      415  -----TVVEGEDVKFTVAKSDS-----KTLDFSDLLTKYNP 447

  QY      279  STGYLYPPYKLVKAADANVGLQYKLNNGNVQVFEFATSTSANNTANTPTPAVDIBIKVAK 338
  DB      448  SV-----SDRISTNTKTN-----TDNHKIAE 468

  QY      339  IVLSGLRFGQ-TIELSVPTCEGNN-----NKVAPMIGNIYLSNENNENADKI 384
  DB      469  ITIKNLKLNOSQTVTLKAKDDSGNVNVEKFTITVQKKEEKQVPKTEQKHSKTEQNVQOE 528

  QY      385  P 385
  DB      529  P 529

```

```

RESULT 14
US-09-268-347-48
; Sequence 48, Application US/09268347
; Patent No. 6335182
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; TITLE OF INVENTION: RECOMBINANT HAEMOPHILUS INFLUENZAE ADHESIN PROTEINS
; FILE REFERENCE: 1038-860
; CURRENT APPLICATION NUMBER: US/09/268,347
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 2048
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-268-347-48

```

Query Match	6.4%;	Score 129;	DB 4;	Length 2048;
Best Local Similarity	21.2%;	Pred. No. 0.49;		
Matches 83;	Conservative 60;	Mismatches 156;	Indels 92;	Gaps
QY	5	ITKKDAN-----PNNGOTQLEAARMELTDLNAKAMTLASQDYAKIEPASIUSAYSEAE	58	
DB	288	TVNNNNRPAYTPN--TQALDPKFOATNNTKAGPLSIGNSIKRKI--INVGAGVYKTD	343	
QY	59	TVNNNLNATLEQLKMAKTNLWESAINQANTDKTTFDNEHPLNVEYKALKTTTLEQRATNLE	118	
DB	344	AVN-----VAQLEAVVWAKERRITFGD-DNSTDVKIGLDNTL-----TIK	384	
QY	119	GLSSTAYNOIRNNLVLYNKA--SLITKTLPDPLNGTLLDLSNLTANKNNINLTSTIN	176	
DB	385	GGAET--NALTDNIGVKEADNSGLVKLAKTLNNLT-----EVNTTTLNATTVKVG	437	
QY	177	EQKTNADALSNSPIKKVIONNEQSFVGTETNANQVPSNYSFVAFSADVTPVNYKYARRTV	236	
DB	438	SSFTTAEILSDSL-----TFTQPN-----	456	
QY	237	WNGDEPSSRIANTNSI--TDVSWIYSLAGTN--TKYQFSFSGWYSGTGYLFPYKLVKA	292	
DB	457	-TGSQSTSKTVYGVNGVKFTNNAAETAAICTTITRDKIGCFARDG--DVDEKQAPYLDKKQ	514	
QY	293	ADANNVCLQYKLVNGNVQVEFATFSANNNTANPTPAYDEIKVAKIVL--SGLRFQGNFI	351	
DB	515	LKVGSVAI--TIDNG-IDAGNKKISNLAKSGSANDAVTIEQLKAARPTLNAGAGISVTPT	571	
QY	352	ELSVPTGEGNNKRVAPMIGNIYLSNNENAD	382	
DB	572	FISVDKASGNV--TAPTY-NIGVKTTFELNSD	599	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```


TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-923-992A-6

Query Match 6.3%; Score 126.5; DB 4; Length 1128;
Best local Similarity 22.5%; Pred. No. 0.33;
Matches 83; Conservative 51; Mismatches 158; Indels 77; Gaps 15;

QY 7 KKDANPNNGOTOLEA-ARMELTDLINA-KAMTLASLODYAKIEASLSAYSE-----AET 59
Db 194 RKQAQADKKREDAEVKREELGKLFSTKAGLDQEIHEHVKKETSSSEENTOKVDEHYANS 253
QY 60 VNNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNL-----VEAYKALKT 108
Db 254 LONLAQKSLEELDRAATN-----EQATQVKNQFLENQAKLKEMQPLIKETNVKLYKAMSE 308
QY 109 TLEQRATNLEGLSTAYNQIRNNLVLYNKASSLITKTLDPNGGTLDDSNEITANKNI 168
Db 309 SLEQVERKELKHNSA-----NLEDLVAKSEIVREYEGKLNQSKNLP--ELKQLEEEA 359
QY 169 NNTLSTINEQ-----KTNADALNSFTKKVI---ONNEQSFVGFTTNANQPSNYSFVAF 220
Db 360 HSKLKQVVEDPRKKFKTSEQVTPKKVRKRDLAANENQOKI-----ELTVSPENI 409
QY 221 SADYTPVNYKYARTVWNGDEPSSRIILANTNSITDVSWIYSLAGTNTKYQFSFS-----N 275
Db 410 -----TVYEGEDVKFTVTAKSDSKTTLDFSLL-----TKYNPSVSDRISTN 451
QY 276 YGPSTG---YLYFPYKLVKAADANNVGLQYKLNNGNQVQVEFATSTSANNTANP-TPAV 331
Db 452 YKTNNDNHKIAEITIKNLKLNESQVTILKAKDDSGNVVKEFTTITVQKKEEKQVPKTPEQ 511
QY 332 DEIKVAKIV 340
Db 512 KDSKTEEKV 520

Search completed: June 12, 2002, 10:49:10
Job time: 203 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:49:47 ; Search time 29.99 Seconds
(without alignments)
1262.394 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086

Perfect score: 2004

Sequence: 1 GCMSTKDKDANPNNGQTLE.....SSNENNADKIPGYRRPGTFL 394

-Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

-Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	831.5	41.5	702	2	S48754	major surface prot
2	754	37.6	702	2	S48753	major surface prot
3	744	37.1	650	2	S48751	major surface prot
4	743	37.1	649	2	S48752	major surface prot
5	713	35.6	647	2	A49218	hemagglutinin homo
6	575.5	28.7	386	2	S48755	major surface prot
7	189.5	9.5	320	2	S51560	major surface prot
8	175	8.7	6713	2	B89921	hypothetical prote
9	173	8.6	1302	1	JG6009	surface-located me
10	165.5	8.3	661	2	A24222	hypothetical prote
11	164.5	8.2	2481	2	D90011	FmtB protein {impo
12	163.5	8.2	3890	2	C89921	hypothetical prote
13	159	7.9	933	2	S41539	fibrinogen-binding
14	156	7.8	807	2	B71605	hypothetical prote
15	155	7.7	1072	2	A68827	hypothetical prote
16	155	7.7	4688	2	F82885	hypothetical prote
17	154.5	7.7	682	2	S44131	subtilisin-like pr
18	154.5	7.7	1365	2	T30822	lmp1 protein - Myc
19	151.5	7.6	1237	2	D71850	probable outer mem
20	151	7.5	624	2	PC6003	surface membrane p
21	150.5	7.5	1051	2	T81351	lmp1 protein - Myc
22	145	7.2	2399	2	D71879	toxin-like outer m
23	144.5	7.2	1487	2	AG2560	hypothetical prote
24	144	7.2	1107	2	AC0976	probable autoirans
25	143.5	7.2	751	2	T40462	ser-lys rich hypot
26	142	7.1	135	2	B49218	hemagglutinin homo
27	142	7.1	2660	2	E85822	probable invasiv z
28	141.5	7.1	1645	2	F36907	phage-related prot
29	140.5	7.0	3194	2	D71917	toxin-like outer m

30 139.5 7.0 2401 2 T28676 rhoptry protein -
31 138.5 6.9 1314 1 TNBYR6 transcription requ
32 137.5 6.9 820 2 T17519 cell surface antiq
33 137.5 6.9 926 2 AE1130 conserved hypothet
34 137.5 6.9 989 2 D89852 fibrinogen-binding
35 137.5 6.9 1524 2 S86553 surface layer prot
36 137.5 6.9 4152 2 T31102 filamentous hemagg
37 137 6.8 719 2 S55119 hypothetical prote
38 137 6.8 5005 2 F82884 hypothetical prote
39 136.5 6.8 568 2 E97066 membrane associate
40 136.5 6.8 1073 2 S14032 kinesin-related pr
41 136.5 6.8 2271 2 F90073 hypothetical prote
42 135.5 6.8 1085 2 T38378 kinesin-like prote
43 135.5 6.8 1238 2 A64596 hypothetical prote
44 135 6.7 2269 2 T28677 rhoptry protein -
45 134.5 6.7 589 2 B97806 hypothetical prote

ALIGNMENTS

RESULT 1

S48754

major surface protein (clone pMGAL.4) precursor - Mycoplasma gallisepticum

C:Species: Mycoplasma gallisepticum

C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999

C:Accession: S48754

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.

FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface

A:Reference number: S48751; MUID:95010739

A:Accession: S48754

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-702 <MAR>

A:Cross-references: EMBL:L28424

C:Genetics:

A:Genetic code: GSG3

A:Start codon: GTG

Query Match 41.5%; Score 831.5; DB 2: Length 702;
Best Local Similarity 42.0%; Pred. No. 3.6e-38;
Matches 173; Conservative 84; Mismatches 116; Indels 39; Gaps 8;

QY 9 DANPNNG-----QTGLEARMELTDLINAKAMTFLASLDYAKIEASLSAYSEATVNN 62
DB 68 NTNPGNGGTTDAAQQQLAAAKKELSDLLATQNSNLSTYADYANIQTTLTAAYTTAKSTD 127
QY 63 NLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSS 122
DB 128 NTSATLEQVKSTSTLTQTAIDTAASKTSFDEKNPELIKAYNAKTKLKKWRNSLSGLTD 187
QY 123 TAYNQRNRLVDLYNKASLLITKTDPLNGGTLDSNEITTANKNNNTLSTNEQKTN 182
DB 188 SNFATIKTNLTALYQSGKDIVTKTLDPLM-GTAINLSAVSQANTNISNAVSKLETWKTN 246
QY 183 DALSNSFIKVTQNNQSFVGTFTNANVOPSNYSVFASFADVTP-----VNYKYARRTV 236
DB 247 TVLATSFVKVEVLVKNLTKGIDT-TNNQEQPGNYSFVGYSVDVTTGSDNARNPNWSFAQRKV 305
QY 237 WNGD-----EPSSRILANTNSITDVSWIYSLAGTNKYQSFNSYGYLYFFPKLVK 291
DB 306 WTSNTDILSQPOPAECENQOSAPDVSWIYNLTGMGAKYSLTFNYGYGSPGFLYFPKLVN 365
QY 292 AADANNVGLQYKLNNGNVQOQVEFATS-----TSANN-----TTANPTPAYD 332
DB 366 SSDSDKVALEYKLINESAVKTIIDFSPSOTSPVASDATRENNRSTAAPAQGSTINPAPTLD 425
QY 333 EIKVAKIVLSGLRFGQNTTSLVPT-TGEGNMKVAIPMIGNIYLSNENNAADK 383
DB 426 DIKIAKVTLSNLKFGNTSTIEFSVPTTAKETSKVAPMIGNMVLTTSSDRDVK 477

RESULT 2

S48753

major surface protein (clone pmGal.3) precursor - Mycoplasma gallisepticum

C:Species: Mycoplasma gallisepticum

C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999

C:Accession: S48753

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;

FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface

A:Reference number: S48751; MUID:95010739

A:Accession: S48753

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-702 <MAR>

A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62417.1; PID:g535690

C:Genetics:

A:Genetic code: SGC3

A:Start codon: GTG

Query Match 37.68; Score 754; DB 2; Length 702;

Best Local Similarity 41.68; Pred. No. 6.2e-34;

Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;

QY 11 NPNNGOT-----OLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVNNLNA 66

Db 69 NPNPNTPEQQLAARKLTLDLCTENTNVALYADYAKIQSTLSAYMTAKTASNTSA 128

QY 67 TLEQLKMAKTNLESAINQANTDKTFDNEHNLVYAKALKTTLEORATNLEGLSSTAYN 126

Db 129 TLENLRASSTTLQAAIDKAANDKRVDSYNOPLVAAYNNLTKLSKTSLSGLSENKYG 188

QY 127 QIRNNLDVLYNKASSLTITKPLDPLNG--GTLDSNEITITANKNINNTLS--TINEQKTN 182

Db 189 GIKNHLSKLFDCGSAITAKTLDPTSGERTLEKVN---ANNKIMAIISPESLKKWKGA 245

QY 183 DALNSFIKKVIONNQSFVGTFTNANQPSNYSFVAFSADVTP-----VNYKYARTVY 237

Db 246 DKF-NEFERNPLSKELKSTSDTAHQEQPANMSFAAYSDLTNSQNLPNWNAQRYW 304

QY 238 NGB-EPSSRILANTN-SITDYSWISYSLAGTNTKYQFSNPGSPGYLYFPYKLYKAAD 294

Db 305 TSENOGPKTALVSSPVSATDVSWSIYSLAGEGTYLTETFGPDNAPLYLPYKLYKAAD 364

QY 295 ANNVGLQYKLNNGNQVQVEP-----ATST-----SAN 321

Db 365 SSSVALQYSLNKTSSKLNFKPAETVSTNTDQSENEVATTSTTEARSSYKVLVADEAATS 424

QY 322 NTTANTPPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNNKVPAMIGNIYLSSENNA 381

Db 425 NNEHNPTVSDINIAKVTLSGLTGGTTEIFSVPEG-----KVAPMIGNIYLSSESQ 479

QY 382 DK1 384

Db 480 VK1 482

RESULT 3

S48751

major surface protein (clone pmGal.1) precursor - Mycoplasma gallisepticum

N:Alternate names: major hemagglutinin pmGA

C:Species: Mycoplasma gallisepticum

C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999

C:Accession: S48751; A44793

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;

FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface

A:Reference number: S48751; MUID:95010739

A:Accession: S48751

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-650 <MAR>

A:Cross-references: GB:U90714; EMBL:L28423; NID:g1905869; PIDN:AAB50152.1; PID:g19058

R:Markham, P.F.; Glew, M.D.; Brandon, M.R.; Walker, I.D.; Whithear, K.G.

Infect. Immun. 60, 3885-3891, 1992

A:Title: Characterization of a major hemagglutinin protein from Mycoplasma gallisepti

A:Reference number: A44793; MUID:92363591

A:Contents: S6

A:Accession: A44793

A:Status: preliminary

A:Molecule type: protein

A:Residues: 26-42 <MA2>

A:Note: sequence extracted from NCBI backbone (NCBIP:1111017)

C:Genetics:

A:Genetic code: SGC3

A:Start codon: GTG

Query Match 37.18; Score 744; DB 2; Length 650;

Best Local Similarity 42.9%; Pred. No. 2e-33;

Matches 172; Conservative 62; Mismatches 127; Indels 40; Gaps 11;

QY 11 NPNNGO-----TOLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEAEVNNN 63

Db 51 NPGDGGQGMNNAQAQELAAARMGLTVFDSRAKNLGLYDYKKTQNTLTAKYDAAKTVLDN 110

QY 64 LNATLEQLKMAKTNLESAINQANTDKTFDNEHNLVYAKALKTTLEORATNLEGLSST 123

Db 111 SSSITQNLNEAKTRLETAIRTAATSKOTFDEQHAELVKVYKELKTTLSNETPATLAPADA 170

QY 124 AYNQIRNNLDVLYNKASSLTITKPLDPLNGGTLDSNETTTANKNINNTL--STINEQKTN 181

Db 171 QYAGIKMHLSDYAGAKAITTKTLEPVGDP-LTACAVTMANTKIVEAIKDEVLNPKKEN 229

QY 182 ADALNSFIKKVIONNQSFVGTFTNANQPSNYSFVAFSADVTP-----NYKYAR 233

Db 230 ATKLADSFVKQVLYKEKITGVEEAHN-RAQPAHYSEFVGYSDVITGTANGQTSIPNNYAO 288

QY 234 RFWW-NGDEPPSRILANT-----NSITDYSWISYSLAGTNTKYQFSNPGSPGYLY 284

Db 289 RTIFTNGDEP--RSVSNTPVDGQTMQAQPLSNVSWISYSLAGTNGAKYLTETFGPSTGYLY 346

QY 285 FPKYKLVKADANNVGLQYKLNNGNQVQVEFATSTSANNTAN-PTPAVDEIKVAKIVLSG 343

Db 347 FPKYKLVNTSDQMKLGLYKLNDA---TEPSAITFGNEQTMNGKTPVNDINNAKVTLAN 402

QY 344 LRFQONTIELSVPTGEGNNKVPAMIGNIYLSSENNAADKI 384

Db 403 LIFGSKNIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 438

RESULT 4

S48752

major surface protein (clone pmGal.2) precursor - Mycoplasma gallisepticum

C:Species: Mycoplasma gallisepticum

C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999

C:Accession: S48752

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.

FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface

A:Reference number: S48751; MUID:95010739

A:Accession: S48752

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-649 <MAR>

A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62416.1; PID:g535689

C:Genetics:

A:Genetic code: SGC3

A:Start codon: GTG

Query Match 37.18; Score 743; DB 2; Length 649;

Best Local Similarity 42.3%; Pred. No. 2.2e-33;

Matches 172; Conservative 62; Mismatches 125; Indels 48; Gaps 11;

```

Qy 9 DANPNNGO-----TQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSSEATVY 61
Db 48 DTNPGDGGMMNAASQELAAARMGLTTFDSKAKNLGLYVDYKKTQNTLTAKYDAKTVL 107
Qy 62 NNLNATLEQLKMAKTNLESAINQANTDKTTEDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db 108 DNSSSTTQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKYKELKTTLSNETATLAPYA 167
Qy 122 STAYNOIRNNLDVLYNKASSLITKTLDPNGTGLDLSNEITTANKNINNTL--STINQK 179
Db 168 DAQYAGIKMHLISGLYDAGKAITTKLEPVEGDP-LTASAVMANTKIIVEATKDEVINPQK 226
Qy 180 TNADALSNSFKKVIQNNQESFVGTFTNANVPQSNYSFVAFSADVPV-----NYKY 231
Db 227 ENATKLADSFVKQVLVKEITGVVEAHN-KAQPANYSFVGSVDITGTTGTSTIPNDY 285
Qy 232 ARRTVW-NGDEPSSRIANT-----NSITDVSWIYSLAGTNKTYQFSNVPSTGY 282
Db 286 AORTFTTNSDEP--RSISNTPADQOTMAQPLSNVSWIYSLAGTGAKYTLEFTYGPSTGY 343
Qy 283 LYFPYKLVKAADANNVGLQYKLNNGNVQOFEATSTSA-----NNTTANPTPAVDKIVA 337
Db 344 LYFPYKLVNTSDQVKGLEYKLN-----ATKPSAITFGSDQTMNGKTPTVNDINVA 395
Qy 338 KIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKI 384
Db 396 KVTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 437

RESULT 5
hemagglutinin homolog pmGAL.2 - Mycoplasma gallisepticum
C:Species: Mycoplasma gallisepticum
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 07-Dec-1999
C:Accession: A49218
R:Markham, P.F.; Glew, M.D.; Whithear, K.G.; Walker, I.D.
Infect. Immun. 61, 903-909, 1993
A:Title: Molecular cloning of a member of the gene family that encodes pmGA, a hemagglut
A:Reference number: A49218; MUID:93162830
A:Accession: A49218
A>Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 1-647 <MAR>
A:Cross-references: GB:955216; NID:g265625; PIDN:AAB25397.1; PID:g265626
A:Experimental source: S6
A:Note: sequence extracted from NCBI backbone (NCBIN:125182, NCBI:P:125183)
C:Genetics:
A:Genetic code: SGC3

Query Match 35.68; Score 713; DB 2; Length 647;
Best Local Similarity 41.18; Pred. No. 9.6e-32;
Matches 169; Conservative 60; Mismatches 124; Indels 58; Gaps 12;

Qy 9 DANPNNGO-----TQLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSSEATVY 61
Db 48 DTNPGDGGMMNAASQELAAARMGLTTFDSKAKNLGLYVDYKKTQNTLTAKYDAKTVL 107
Qy 62 NNLNATLEQLKMAKTNLESAINQANTDKTTEDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
Db 108 DNSSSTTQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKYKELKTTLSNETATLAPYA 167
Qy 122 STAYNOIRNNLDVLYNKASSLITKTLDPNGTGLDLSNEITTANKNINNTL--STINQK 179
Db 168 DAQYAGIKMHLISGLYDAGKAITTKLEPVEGDP-LTASAVMANTKIIVEATKDEVINPQK 226
Qy 180 TN-----ADALSNSFKKVIQNNQESFVGTFTNANVPQSNYSFVAFSADVPV----- 227
Db 227 ENATKLADSLSSIVKKITGVVE-----AHNKAQPANYSFVGVKRYWTELLDKQVFP 279
Qy 228 NYKYARRTVW-NGDEPSSRIANT-----NSITDVSWIYSLAGTNKTYQFSNVPSTGY 278
Db 280 NWDYQAORTIFTNSDEP--RSISNTPADQOTMAQPLSNVSWIYSLAGTGAKYTLEFTYGP 337

```

```

Qy 279 STGYLYFPYKLVKAADANNVGLQYKLNNGNVQOFEATSTSA-----NNTTANPTPAVD 333
Db 338 STGYLYFPYKLVNTSDQVKGLEYKLN-----ATKPSAITFGSDQTMNGKTPTVND 389
Qy 334 IKVAKIVLSGLRFGQNTIELSVPTGEGNMKNVAPMIGNIYLSNENNADKI 384
Db 390 INVAKVTTLANLNFSGNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNKI 435

RESULT 6
major surface protein (clone pmGAL.5) precursor - Mycoplasma gallisepticum (fragment)
S48755
C:Species: Mycoplasma gallisepticum
C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S48755
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surfac
A:Reference number: S48751; MUID:95010739
A:Accession: S48755
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-386 <MAR>
A:Cross-references: EMBL:L28424; NID:g535687; PIDN:AAA62419.1; PID:g535692
A:Note: the sequence of residues 385-386 and the corresponding nucleotide sequence ar
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match 28.78; Score 575.5; DB 2; Length 386;
Best Local Similarity 42.38; Pred. No. 1.6e-24;
Matches 138; Conservative 50; Mismatches 113; Indels 25; Gaps 10;

Qy 9 DANPNNG-----QTOLEAARMELTDLINAKAMTSLASQDYAKIEASLSAYSSEATVY 62
Db 64 NTNPGCGGMDNSAQALAAAKKELSDLLATQNSLSTVADYAKIKNDLTAAYTTAETASQ 123
Qy 63 NNLNATLEQLKMAKTNLESAINQANTDKTTEDNEHPNLVEAYKALKTTLEQRATNLEGLS 122
Db 124 NQAATLEQVKNAASTLQTAINTAVNEKVFDENNSELSTAYTNLTKTLEGENTLAAAFND 183
Qy 123 TA-YNOIRNNLDVLYNKASSLITKTLDPNL-GTLLDSNEITTANKNINNTL--STINQK 178
Db 184 SANYGGIKTHLLSLYNQAKTITTTSL--LNDAGQSPKNDVNVKINKEITDAINPTLLNQ 241
Qy 179 KTNADALSNSFKKVIQNNQESFVG---TFTNANVPQSNYSFVAFSADVP-----VNY 229
Db 242 KANADMLATSFTKQVL-NDAQLTSGSSETSMOTQPGNYSFVGYSVDVTTGSNNARPNN 300
Qy 230 KYARTVWNGDEPSSRIANTNSITDVSWIYSLAGTNKTYQFSNVPSTGY-TCYLYFPYK 288
Db 301 NFAQRKVVDTNRAPLAQTEQSNKLDVSWIYSLSGMGAKYTTVFDYYGASNNAYLYFPYK 360
Qy 289 LVKAADANNVGLQYKLNNGNVQOFEF 314
Db 361 LVQVND--NVGLQYVNLNTPPKLVNF 384

RESULT 7
major surface protein (clone pmGAL.6) - Mycoplasma gallisepticum (fragment)
S51560
C:Species: Mycoplasma gallisepticum
C:Date: 01-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
C:Accession: S51560; S48757
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.
FEBS Lett. 352, 347-352, 1994
A:Title: The organisation of the multigene family which encodes the major cell surfac
A:Reference number: S48751; MUID:95010739
A:Accession: S51560
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA

```



```

Qy 9 DANPNNGOTQLE-----AARMELTDLINAKAMTILASIQDYAKIAEIASLSSAYSEAE 58
Db 1186 DOAKTTGETSIDVPTVVKKATAREITAILUNLQEIQTDPDATDEEKQADA--EAN 1243
Qy 59 TVNNLN-----ATLEQLKMAKTLESAINOANTDKTTFDNEHPLNVLVAYKALKTTL-- 110
Db 1244 TENGKANQAI SAATTNAQVDEAKANAERAIN-AVTPKVVKQAAKDEIDOLQATQTNVIN 1302
Qy 111 -EQRATNLEGLS-----STAYNQIRNNL-----VDLYNKASSLIITKTLDPNGGTLL 156
Db 1303 NDQONATNEEKAAIQOLATAVTADAKNNITAAATDDNGVDVTAKDAGKNSIQSTOP--ATAV 1359

```

Db 1360 KSNAKNEVDQAVTTNQALDNTGATTEKNAKDL-----VLKAKEKAYQDIL---N 1409

Qy 210 VQPSNYSFVAFSADYTPVNYKARYRTVWNGDEPSSRILANTSITDVSWYSLA-----GT 265

Db 1410 AQTN-----DVTQI-----KDAQVADIOGTTADTTIKDVAKDELAT 1446

Qy 266 NTKYQFSPSNYCPSTGYLYFPKLYLKAADAN-----NVGLQYKLNNNGNVQVEFATST 318

Db 1447 KANEQKAL-----IAQTADATTEKEQANQOQVDAQUTQGN-QNIEAQS 1490

Qy 319 SANNTTA-NPTPAVDKIVAKIVLSGLRPGQNTIELSVPTGEGNMKNVAPMIGNIVLSSN 377

Db 1491 DDVNTAKDNAIQADIPQIAQSTDVKYNAR-----AELLTEM-----QNKITIELNNNETTNE 1541

Qy 378 ENNAD 382

Db 1542 EKGND 1546

RESULT 12

C89921

hypothetical protein ebhB [Imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: C89921

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui
ma, A.; Mizutani-Uh, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Seki
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146

```

A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-3890 <KUR>
A: Cross-references: GB:BA000018; PID: g13701233; PIDN: BAB42528.1; GSPDB: GN00149
A: Experimental source: strain N315
C: Genetics:
A: Gene: ebhB

Query Match      8.2%; Score 163.5; DB 2; Length 3890;
Best Local Similarity 24.4%; Pred. No. 1.1;
Matches 94; Conservative 54; Mismatches 153; Indels 85; Gaps 20

Qy 17 TLEAARMELTDLIINAKMTLASLDYAKIEASLSAYSEA-ETVNNLNMTLEQLKMAK 75
Db 3528 TDLNTAMGNLQAINDEQPTLLS-QNYODATPSKKTAYTNAVOAAKDILNKSQGNK-TK 3585
Qy 76 TNLSEAINQANTDKTTFDNEHENLVEAYKALKTTLEQRATNLGLSSTAYNQIRNNLVDL 135
Db 3586 DQVTEAMNQVNSAKNNLDG-----TRLLOPAOKTAKQQLNNMTHLTQAQKTNLTNQ---- 3636

```

Db	3637	-----INSGTTVAGVHTVQSNANTLDQAMNPLRQSTANNDATKASDEYVDA	3682
QY	191	----KVIQNNQESVGTETNANVOPS-NYSEFAFSADVTVPVNYKYARRTVWNGDE----	241


```

Db 430 ESNVSQDTSSEAS-TNSNSSISLSPSNISSTSD--SESATNSSDFSNVAEVANNLSASV 486
QY 65 -NATLEOLKMAKTNLESALNCAQNTDKTTFDNEHNLVYAYKALKTTLEQRATN----- 116
Db 487 NNSSSVLSSTSTADNLGINOSGSDNLTKDSSEISTSGAFLSSNOTSSEASTNSNISL 546
QY 117 -----LEG-LSSTAYNQI----RNNLVDLYNKASSLI--TKTLDPL-----NGG 153
Db 547 SPSNISSTVLESTSSSNFNSVAEVANNLSASVYNNSSSVLSSTSTADNLEINQFGSDN 606
QY 154 TLLDSNEITTANKNINNTLSTINEOKTNADALSNSFTKKVYIONNEQSFVGTFTNANVOPS 213
Db 607 LTKDSEISTSG-----AFLSSNOTSSEASSNS-----MSSINSPSLSLSTNSSESAT 655
QY 214 NYSFVAFSADVTPYNYKYARTVNGDEPSSRII-----ANTNSITDVSWIYSLAG 264
Db 656 NQS---NSSEATKYDN-----NSSTHSSNIISSGSDSDSDSDSSNLSLSSPN 704
QY 265 TNTKYQFSFSGNYGPGTGYLPYKLVKAADANNYGLOYKLNGNVQOQVEFATSTSANNT- 323
Db 705 LETNOTIS-----SKPSEVNNI-----SENPKVSSNSVQENSTD 740
QY 324 ---TANTPAVDE-----IKVAKIVLGLRPFQONTIE---LSVPTGEG 360
Db 741 HEMSTNPKSSISPISTTSSSQKESQSNLLTTTEGINNPITFNSSSENSEAAASILATYS 800
QY 361 MNKVAPMIGNIYLS-----SNNENADKI 384
Db 801 NNSSESSETGCLYISNEAQRDNGSEISHSLPSSNENNVSSI 843

```

Search completed: June 12, 2002, 10:49:50
Job time: 208 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:51:05 ; Search time 17.44 seconds
(without alignments)
874.742 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086

Perfect score: 2004

Sequence: 1.GCMSTTKDANPNNGQTGLE.....SSNENADKIPGYRREGTEL 394

*Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

*Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	154.5	7.7	682	1 NISP_LACLA	Q07596 lactococcus
2	143	7.1	857	1 AR56-CANAL	P78586 c arg5,6 pr
3	138.5	6.9	1314	1 SWI1_YEAST	P09547 saccharomyc
4	137	6.8	719	1 YM41_YEAST	Q03213 saccharomyc
5	135.5	6.8	1085	1 CUT7_SCHPO	P24339 schizosacch
6	132	6.6	1271	1 Y338_MYCGE	P47580 mycoplasma
7	131.5	6.6	1164	1 BRG1_STRAG	P27951 streptococc
8	129	6.4	1251	1 RBP2_PLAVB	Q00799 plasmodium
9	128.5	6.4	2334	1 WAPA_BAGSU	Q07833 bacillus su
10	128.5	6.4	3712	1 LMA_DROME	Q00174 drosophila
11	128	6.4	995	1 YI09_YEAST	P40442 saccharomyc
12	127	6.3	956	1 YEF3_YEAST	P32618 saccharomyc
13	126.5	6.3	444	1 PST1_YEAST	Q12355 saccharomyc
14	126.5	6.3	1630	1 MSP1_PLAFK	P04932 plasmodium
15	126.5	6.3	1639	1 MSP1_PLAFW	P04933 plasmodium
16	125.5	6.3	1858	1 P3K2_DICDI	P54674 dictyosteli
17	124.5	6.2	821	1 LIN5_CABEL	P45970 caenorhabdi
18	124.5	6.2	1010	1 YAK1_CABEL	P34278 caenorhabdi
19	124.5	6.2	1018	1 FNBA_STAAU	P14738 staphylococ
20	124	6.2	749	1 YKSI_YEAST	P40957 saccharomyc
21	124	6.2	750	1 YKSI_YEAST	P34231 saccharomyc
22	124	6.2	823	1 NSP1_YEAST	P14907 saccharomyc
23	124	6.2	866	1 MYSP_SCHJA	Q05870 schistosoma
24	124	6.2	1117	1 YN96_YEAST	P33753 saccharomyc
25	124	6.2	1159	1 N124_SCHPO	Q09904 schizosacch
26	124	6.2	1433	1 CAT8_YEAST	P39113 saccharomyc
27	123.5	6.2	537	1 ARP_PLAFA	P04931 plasmodium
28	123.5	6.2	1902	1 P2P_LACLC	P15293 lactococcus
29	123	6.1	1260	1 ALS1_CANAL	P46590 candida alb
30	123	6.1	1637	1 MKSP_STAAU	P80544 staphylococ
31	123	6.1	1790	1 USO1_YEAST	P25386 saccharomyc
32	122.5	6.1	507	1 FLIC_SALON	Q06974 salmonella
33	122	6.1	650	1 PDR4_YEAST	P19880 saccharomyc

RESULT 1

NISP_LACLA

ID	NISP_LACLA	STANDARD;	PRT;	682 AA.
AC	Q07596;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Nisin leader peptide processing serine protease NISP precursor			
DE	(EC 3.4.21.-).			
GN	NISP.			
OS	Lactococcus lactis (subsp. lactis) (Streptococcus lactis).			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Lactococcus.			
OX	NCBI_TaxID=1360;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN=NIZO R5;			
RX	MEDLINE=93239683; PubMed=8478324;			
RA	van der Meer J.R., Polman J., Beerthuyzen M.M., Siezen R.J.,			
RA	Kuipers O.P., de Vos W.M.;			
RT	"Characterization of the Lactococcus lactis nisin A operon genes			
RT	nisp, encoding a subtilisin-like serine protease involved in			
RT	precursor processing, and nispR, encoding a regulatory protein			
RT	involved in nisin biosynthesis."			
RT	J. Bacteriol. 175:2578-2588(1993).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=6F3;			
RX	MEDLINE=94213458; PubMed=8161176;			
RA	Engelke G., Gutowski-Eckel Z., Kiesau P., Siegers K.,			
RA	Hammelmann M., Entian K.-D.;			
RT	"Regulation of nisin biosynthesis and immunity in Lactococcus lactis			
RT	6F3."			
RL	Appl. Environ. Microbiol. 60:814-825(1994).			
RN	[3]			
RC	3D-STRUCTURE MODELING.			
RX	MEDLINE=95357326; PubMed=7630881;			
RP	Siezen R.J., Rollemans H.S., Kuipers O.P., de Vos W.M.;			
RA	"Homology modelling of the Lactococcus lactis leader peptidase Nisp			
RT	and its interaction with the precursor of the lantibiotic nisin."			
RT	Protein Eng. 8:117-125(1995).			
CC	-1- FUNCTION: CLEAVES THE LANTIBIOTIC NISIN PRECURSOR PEPTIDE.			
CC	-1- PATHWAY: LAST STEP OF NISIN BIOSYNTHESIS.			
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL			
CC	(POTENTIAL).			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE			
CC	SUBTILASE FAMILY.			
CC	-1- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS			
CC	IN THE REGION OF THE MEMBRANE ANCHOR.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/)			

34	122	6.1	1672	1 PMPB_CHLMU	Q09jy2 chlamydia m
35	121.5	6.1	425	1 YBY0_YEAST	P38272 saccharomyc
36	121.5	6.1	1140	1 YM96_YEAST	Q04893 saccharomyc
37	121	6.0	491	1 TOLC_SALEN	Q54001 salmonella
38	121	6.0	796	1 YS8A_CAEEL	Q09625 caenorhabdi
39	121	6.0	901	1 PIP_LACLA	P49022 lactococcus
40	121	6.0	979	1 P115_MYCHR	P41508 mycoplasma
41	121	6.0	1111	1 KIP1_YEAST	P28742 saccharomyc
42	121	6.0	1902	1 P2P_LACPA	Q02470 lactobacill
43	120.5	6.0	776	1 VP4_ROTCH	Q02945 equine rota
44	120.5	6.0	800	1 INLA_LISMO	P25146 listeria mo
45	120.5	6.0	1565	1 PAC_STRMU	P11657 streptococc

ALIGNMENTS

```

or send an email to license@isb-sib.ch).
-----
EMBL: L11061; AAA25200.1; -
EMBL: X76884; CAA54210.1; -
HSP: P29600; IGI.
MEROPS: S08.059; -
InterPro: IPR001899; Gram_pos_anchor.
InterPro: IPR002029; Peptidase_S8.
Pfam: PF00082; Peptidase_S8; 1.
PRINTS: PR00723; SUBTILISIN.
PROSITE: PS00136; SUBTILASE_ASP; 1.
PROSITE: PS00137; SUBTILASE_HIS; 1.
PROSITE: PS00138; SUBTILASE_SER; FALSE_NEG.
PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
KW Hydrolyase; Serine protease; Cell wall; Signal; Transmembrane; Zymogen.
FT SIGNAL 1 22
FT PROPEP 23 195
FT CHAIN 196 682
FT DOMAIN 196 651
FT TRANSMEM 652 679
FT DOMAIN 680 682
FT ACT_SITE 259 306
FT ACT_SITE 306 306
FT ACT_SITE 512 512
FT DOMAIN 652 657
FT CONFLICT 500 500
FT SEQUENCE 682 AA; 74767 MW; D5F29313F2983EC9 CRC64;

Query Match 7.7%; Score 154.5; DB 1; Length 682;
Best Local Similarity 21.0%; Pred. No. 0.11;
Matches 75; Conservative 61; Mismatches 130; Indels 91; Gaps 13;

QY 58 ETVNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTKTLEQRTNL 117
DB 38 ELINHNNAISLSTEGSTDSINLGAQSPAVKSTTRTE-----LDVTGAAKTLTQTSVQK 93
QY 118 EGLSTAYNQIRNLDVYNKASLTKTLDPLNGTGLDLSNIEITANKNNINNTLTINE 177
DB 94 EMKVSQETQVSSE-----FSKRDSVTNKEAVPVSKDELLEQSEVVSTSIQKN-KILDN 148
QY 178 QKTNADALSFTKVKVQNNQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARTV- 236
DB 149 KKKRANFTVSPLIKKPSNKSQASGVIDNS-----ASPLSYKRAKEVWS 193
QY 237 -----WNGDEPSSRLTANSTIDVSWIYSLAGTNTKYQFSFNSYGPS 279
DB 194 LRQPLKQKVEAOPLLISNSEKKASVYTNSHDFWYQW-----DMKY---VTNNGES 243
QY 280 TGYLYFPYKLVKAADANNVGLQYKLNNGVQVQVEFATSTAN-----NT 323
DB 244 YA-LYQPSKKI-----SVGI---IDSGIMEHPDLSNLGNFYKFLVPKGGFDNEERPDE 293
QY 324 TANTPVADEIKVAKVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIG-NIYLSNEN 379
DB 294 TGNPSDIVD-----KMGHGTEVAGQITANGNILGVAPEGITVNIYRVEGEN 338

RESULT 2
AR56_CANAL STANDARD; PRT; 857 AA.
AC P78586;
DT 01-NOV-1997 (Rel. 35, Created)
DE 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE ARG5,6 protein, mitochondrial precursor [Contains: N-acetyl-gamma-
DE glutamyl-phosphate reductase (EC 1.2.1.38) (N-acetyl-glutamate
DE semialdehyde dehydrogenase) (NAGSA dehydrogenase); Acetylglutamate
DE kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-glutamate 5-
DE phosphotransferase)].
GN ARG5,6.

```

```

OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 64385 / 1001;
RX MEDLINE=97195775; PubMed=9043106;
RA Negredo A., Monteoliva L., Gil C., pla J., Nombela C.;
RT "Cloning, analysis and one-step disruption of the ARG5,6 gene of
RT Candida albicans.";
RL Microbiology 143:297-302(1997).
CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
CC + phosphate -> N-acetyl-5-glutamyl phosphate + NADPH.
CC -1- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate -> ADP + N-acetyl-L-
CC glutamate 5-phosphate.
CC -1- PATHWAY: SECOND AND THIRD STEPS IN ARGININE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC ACETYLGLUTAMATE KINASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAGSA
CC DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X98880; CAA67383.1; -
CC InterPro: IPR000706; AGPR_act_site.
CC InterPro: IPR001048; Akinase.
CC InterPro: IPR000534; Semiadh_dh.
CC Pfam: PF00696; aakinase; 1.
CC Pfam: PF01118; Semialdehyde_dhc; 1.
CC Pfam: PF02774; Semialdehyde_dhc; 1.
CC ProDom: PD003765; AGPR_act_site; 1.
CC PROSITE: PS01224; ARG; 1.
CC Oxidoreductase; transferase; Kinase; Arginine biosynthesis; NADP;
CC Mitochondrion; Multifunctional enzyme; transit peptide.
CC TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
CC CHAIN ? ? ACETYLGLUTAMATE KINASE.
CC CHAIN ? 857 N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE
CC REDUCTASE.
CC ACT_SITE 669 669 BY SIMILARITY.
CC SEQUENCE 857 AA; 95066 MW; AE2CEAD8FF8C471 CRC64;

Query Match 7.1%; Score 143; DB 1; Length 857;
Best Local Similarity 23.3%; Pred. No. 0.64;
Matches 99; Conservative 58; Mismatches 134; Indels 134; Gaps 25;

QY 46 IEASLSAY-----SRAETVNN---NLNATL-----EQLKMAKTNLESAINQANTD 88
DB 208 VEAAINSGLPILTSALTSSGQLLNADVAAAGELAREFEPLKIVYLNKGGIINGTG 267
QY 89 K-----TTFDNEHPNLV-EAYKALKTKTLEQR-----ATNLEGLSSTAYNQIRNLDVLYNK 138
DB 268 EKVSAINLDEEYEDLLKESWVYGTGKIKIHDLOHLPRSSVA-----IIDVNDL 320
QY 139 ASSLIITKTLPLNGTGLDLSNIEITANKNNIN-NTLSTINEQKTNADALSFTKVKVQNN 197
DB 321 QKELFTDS-----GAGTL-----IRRGYRLNRLNSLRDFG---NPDLRLNALLR----- 361
QY 198 EQSFVGFTTANVQPSNYSFVAFSADVTPVNYKYARTVWNGDEPSSRIANT----- 250
DB 362 -----DPEIKTGKVSVAAYLKFLDSVQFKS-----YGDPE-LEVLAIVVEQNDKI 405
QY 251 -----NSITD-----VSWIYSLAGTNTKYQFSFNSYNG-PSTGYL 283
DB 406 PKLDEFSSKGTGLNNVTDITFNIFNAIKKDYSLQCVWVNDANILPWYFSKDSGSAKNGQI 465

```

```

Query Match      6.8%; Score 137; DB 1; Length 719;
Best Local Similarity 19.8%; Pred. No. 1.1;
Matches 105; Conservative 72; Mismatches 175; Indels 178; Gaps 23;

Qy 3 MSITKIDANPNNGQTQLEAARMETDL-----INAKMTLAS----- 39
Db 38 MPTTLKDCYRNNSOVNEDAIGINL-DLSLPTHTSPPTGSESAGSGNATLRNDGNALDGG 96

```


Query Match		6.6%; Score 131.5; DB 1; Length 1164;	
Best Local Similarity		22.8%; Pred. No. 4;	
Matches		84; Conservative	51; Mismatches 157; Indels 77; Gaps 15;
QY	7	KKDANPNNGQTOLFA-ARMELTDLINA-KAMTLASLDQYAKIEASLSAYSE-----AET	59
DB	230	RKQAQADKKEDAEVKVREELGKLFSTKAGLDQEIQEHVKKETSSEENTQKVDEHYANS	289
QY	60	VNNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNL-----VEAYKALKT	108
DB	290	LONLAQKSELELDKATN-----EQATQVKNOFLNAQKLKEIQIOLIKETNVKLYKAMSE	344
QY	109	TLQERATNLEGLSTPAYNOIRNNLVLYNKASSLITKTLDPINGTLLDSNEITANKNI	168
DB	345	SLEQVEKELKHNSA-----NLEDLVAKSEIVREYEGKLNQSNLP--ELKOLEEEA	395
QY	169	NNTLSTINEQ-----KTNADALSNSFIKKVI---QNNQSFVGTFTNANVOPSNYSFVAF	220
DB	396	HSKLQGVVEDFRKKFKTSQVTPPKRVARDLAANNENNOQKI-----ELTVSPENI	445
QY	221	SADVTPVNYKYARRTVWNGDEPSSRILANTNSITDVSIIYSLAGTNTKYQFSFS-----N	275
DB	446	-----IVYEGEDVKFTVTAKSQSKTTLDFSDLL-----TKYNPSVSDRISTN	487
QY	276	YGPSTG-----YLYFPYKLVKAADANNVGLQYKLNNGNVQVFEATSTSANNTTAND-TPAV	331
DB	488	YKTTNDNHKIAEITIKNLKLNESQTVTLKAKDDSGNVVEKFTITTVQKKEEQVPTPEQ	547
QY	332	DEIKYAKIV	340
DB	548	KDSKTEERV	556
RESULT 8			
ID	RBP2_PLAVB	STANDARD;	PRT; 1251 AA.
AC	Q00799;		
DT	01-APR-1993 (Rel. 25, Created)		
DT	01-APR-1993 (Rel. 25, Last sequence update)		
DT	01-OCT-1996 (Rel. 34, Last annotation update)		
DE	Reticulocyte binding protein 2 (Fragment).		
GN	RBP2.		
OS	Plasmodium vivax (strain Belem).		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_TaxID=31273;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=92315338; PubMed=1617731;		
RA	Gallinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;		
RT	"A reticulocyte-binding protein complex of Plasmodium vivax		
RT	merozoites."		
RL	Cell 69:1213-1226(1992).		
CC	-1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO		
CC	HUMAN RETICULOCYTE CELLS.		
CC	-1- SUBCELLULAR LOCATION: Membrane-bound (Probable).		
CC			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	at the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC			
DR	EMBL; M88096; AAA29744.1; -		
KW	Malaria; Receptor; Membrane.		
FT	NON_TER	1	
FT	NON_TER	1251	1251
SEQ	SEQUENCE	1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;	
Query Match			
Best Local Similarity		6.4%; Score 129; DB 1; Length 1251;	
Matches		21.1%; Pred. No. 6;	

Matches		89; Conservative	70; Mismatches 140; Indels 122; Gaps 20;
QY	16	QTOLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEAE-----TVNNLNAT	67
DB	424	ETNLESYKHNLSITNIK-----QGGEKI-----YSKADIMQIKATSENTAETK	469
QY	68	LEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSSTAYNQ	127
DB	470	LEKVKDDQSNVYVNLQITTER-----NLIVTERNLNGIDSTITNIEG-----ALKE	517
QY	128	IRNNL-----VDLYNKASSLITKTLDPINGTLLDSNEIT--TANKN	167
DB	518	SKGNYEIGFLEKEEIGKRRKLVKITKS---INSTVG--NFSSLFNFDLNOYDFNKN	572
QY	168	INNTLSITNEQKTNADALSNSF---IKVQNNQSFVGTFTNANVOPSNYSFVAFSADY	224
DB	573	INDYENKGE-----IYNEFEGLSKISEN-----LRNASENTSDYN---SAKT	613
QY	225	TPVNYKYARRTVWNGDEPSSRILANTNSITDVSIIY-----SLAGTN---TKYQFSFSNYG	277
DB	614	LRLEAQREKVNLLNKEEANKYLRDVKVFSFRFTFNKESLDKINEMIKKEQLTVNEGH	673
QY	278	PSTGYLYPPYKLVKAADANNVGLQYKLNNGNVQVFEATSTSANN-----TTA--	325
DB	674	GNVQQLVENIK--ELVDENNLSDLKQATGKNEEQKITHSTLKNKAKTILGHVDTFSKY	731
QY	326	-----NPTPAYDE-----IKVAK-----IVLSGLRFGQNTIELSVPTGEGNMKV	365
DB	732	VGKITPELALTELLGDAKLKTAQELKFESKNVNVLETNENSKNTNELDVHKNIDQAVKV	791
QY	366	A	366
DB	792	A	792
RESULT 9			
ID	WAPA_BACSU	STANDARD;	PRT; 2334 AA.
AC	O07833;		
DT	01-OCT-1994 (Rel. 30, Created)		
DT	01-OCT-1994 (Rel. 30, Last sequence update)		
DT	01-MAR-2002 (Rel. 41, Last annotation update)		
DE	Wall-associated protein precursor.		
GN	WAPA OR NI7G.		
OS	Bacillus subtilis.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;		
OC	Bacillus/Staphylococcus group; Bacillus.		
OX	NCBI_TaxID=1423;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=93302506; PubMed=8316082;		
RA	Poster S.J.;		
RT	"Molecular analysis of three major wall-associated proteins of		
RT	Bacillus subtilis 168: evidence for processing of the product of a		
RT	gene encoding a 258 kDa precursor two-domain ligand-binding		
RT	protein."		
RL	Mol. Microbiol. 8:299-310(1993).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	STRAIN=168 / BGSC1A1;		
RX	MEDLINE=95219088; PubMed=7704263;		
RA	Yoshida K.-I., Sano H., Seki S., Oda M., Fujimura M., Fujita Y.;		
RT	"Cloning and sequencing of a 29 kb region of the Bacillus subtilis		
RT	genome containing the hut and wpa loci."		
RL	Microbiology 141:337-343(1995).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RX	STRAIN=168 / BGSC1A1;		
RX	MEDLINE=97124196; PubMed=8969509;		
RA	Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,		
RA	Miwa Y., Fujita Y.;		
RT	"Sequencing of a 65 kb region of the Bacillus subtilis genome		

containing the lic and cel loci, and creation of a 177 kb contig covering the gnt-sacxy region.";

Microbiology 142:3113-3123(1996).

-1- FUNCTION: STILL UNKNOWN. NOT INVOLVED IN CELL MEMBRANE METABOLISM, MOTILITY, SECRETION OR DIFFERENTIATION.

-1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MAY BE RELEASED INTO THE MEDIUM.

-1- DOMAIN: HAS TWO LIGAND-BINDING DOMAINS; THE N-TERMINUS, HAS THREE 101 AA REPEATS WHICH ARE RESPONSIBLE FOR CELL WALL BINDING; THE C-TERMINUS CONSISTS OF TWO BLOCKS OF RESIDUES WITH A CONSERVED MOTIF REPEATED 31 TIMES.

-1- SIMILARITY: THE REPEATED MOTIF IN THE C-TERMINUS HAS SOME SIMILARITY TO THE REPEAT IN E. COLI RHS GROUP OF PROTEINS (RHS-A-D).

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL: L05634; AAA22883.1; -
 EMBL: D31856; BAA06556.1; -
 EMBL: D29985; BAA06260.1; -
 EMBL: D83026; BAA11683.1; -
 EMBL: Z99124; CAB15959.1; -
 PIR: S32920; S32920.
 Subtilist; BGL0797; wapa.
 InterPro: IPR003305; CBD_6.
 Pfam: PF02018; CBD_6; 1.
 Cell wall; Repeat; Signal; Complete proteome.
 SIGNAL 1 28
 FT CHAIN 29 2334
 FT DOMAIN 504 869
 FT REPEAT 504 605
 FT REPEAT 636 736
 FT REPEAT 769 869
 FT DOMAIN 1021 2139
 FT REPEAT 1021 1040
 FT REPEAT 1042 1061
 FT REPEAT 1063 1082
 FT REPEAT 1083 1102
 FT REPEAT 1109 1128
 FT REPEAT 1129 1148
 FT REPEAT 1150 1169
 FT REPEAT 1174 1193
 FT REPEAT 1199 1218
 FT REPEAT 1219 1238
 FT REPEAT 1646 1665
 FT REPEAT 1667 1686
 FT REPEAT 1690 1709
 FT REPEAT 1711 1730
 FT REPEAT 1732 1751
 FT REPEAT 1753 1772
 FT REPEAT 1795 1814
 FT REPEAT 1820 1839
 FT REPEAT 1840 1859
 FT REPEAT 1861 1880
 FT REPEAT 1887 1906
 FT REPEAT 1908 1927
 FT REPEAT 1929 1948
 FT REPEAT 1969 1982
 FT REPEAT 1983 2002
 FT REPEAT 2008 2027
 FT REPEAT 2028 2047
 FT REPEAT 2051 2070
 FT REPEAT 2071 2090
 FT REPEAT 2093 2112
 FT REPEAT 2120 2139
 FT REPEAT 2334 AA; 258329 MW; B75138CCD278BAA3 CRC64;
 SQ SEQUENCE

Query Match 6.4%; Score 128.5; DB 1; Length 2334;
 Best Local Similarity 22.1%; Pred. No. 14;
 Matches 91; Conservative 56; Mismatches 152; Indels 113; Gaps 22;

Qy 4 SITKDDANPNQCQLEARM-ELTDLINAKAMTULASLDYAKIASLSAYSEATVN- 61
 Db 1772 SVANKEQNTTKRTPDNKRNLTELTDRGSGQTWTPSDSKLK---TFSWIHGDKGTNQ 1828
 Qy 62 ---NNLNATLQOLMAKT-----NLESAINQANTDKTTFEDNEHNLV----- 100
 Db 1829 FTYNKLOMIEKMDSTSYSDYDENGWVQTFIT-GNGGTSFSYDERNLVSSLHIGDKN 1887
 Qy 101 -----EAYK-----ALKTTLEQRATNLEGLSTAYNQIRNNLVLYNKASSLITKTLDP 150
 Db 1888 GGDILTSEYDANGNRRTINSSAS---GKQVYEGKL-NQLVKETHEDGTVEITYDGF 1943
 Qy 151 NG-----GYLLD-SNEITANKNNNTLTSTNEOKTNADALSNSFKKVIQNEQSFVGTFF 205
 Db 1944 GNRKTVTTIKDGSSTVNASFNIMNQLTKVNDESISYD-----KNGNRRTSDGKF 1992
 Qy 206 TNANVQPSNYSFVAFSADVTP-VNYKYARRTVWNGDEPSSRLANTNSITDVSWLYSLAG 264
 Db 1993 TYTWDADNLTAVTKKGDKFPATYK-----DEGNRIQKTVNGKV-TNYFYDGD 2043
 Qy 265 TMTKYQFSFN-----YGPS-----TCYLYFPYKLVKAADANNVGLQYKLN-NGN 308
 Db 2044 LNVLYETDADNNVTKSYTYGDSGQLLSYENGKKYF-----YHNAHGD 2087
 Qy 309 VOQVEFATSTSANNTA-----NP--TPAVDEIKVAKIVLSGLRFGQNT 350
 Db 2088 I-----IAISDSTGKTVAKYQYDAMGNPTKTEASDEVKONRYAGYQYDEET 2135

RESULT 10

LMA_DROME
 ID LMA_DROME STANDARD; PRT; 3712 AA.
 AC Q00174;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 01-MAR-2002 (Rel. 41, Last annotation update)
 DT Laminin alpha chain precursor.
 GN LANA OR LAMA.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93049203; PubMed=1425586;
 RA Kusche-Gullberg M., Garrison K., Mackrell A.J., Fessler L.I.,
 RA Fessler J.H.;
 RA "Laminin A chain: expression during Drosophila development and
 RT genomic sequence.";
 RL EMBO J. 11:4519-4527(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=94038678; PubMed=8223265;
 RA Henchcliffe C., Garcia-Alonso L., Tang J., Goodman C.S.;
 RA "Genetic analysis of laminin A reveals diverse functions during
 RT morphogenesis in Drosophila.";
 RL Development 118:325-337(1993).
 RN [3]
 RP SEQUENCE OF 1762-3712 FROM N.A.
 RX MEDLINE=92078147; PubMed=1744083;
 RA Garrison K., Mackrell A.J., Fessler J.H.;
 RA "Drosophila laminin A chain sequence, interspecies comparison, and
 RT domain structure of a major carboxyl portion.";
 RL J. Biol. Chem. 266:22899-22904(1991).
 CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
 IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF

CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
 CC -|- FUNCTION: DIVERSE FUNCTIONS DURING MORPHOGENESIS IN DROSOPHILA.
 CC COMPLETE LOSS-OF-FUNCTION MUTATIONS LEAD TO LATE EMBRYONIC
 CC LETHALITY. CERTAIN PARTIAL LOSS-OF-FUNCTION MUTATIONS GIVE RAISE
 CC TO ESCAPER ADULTS, WHICH HAVE ROUGH EYES ASSOCIATED WITH CHANGES
 CC IN CELL FATE AND PATTERN, MISSHAPEN LEGS AND DEFECTS IN WING
 CC STRUCTURE.
 CC -|- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
 CC -|- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
 CC MEMBRANES (MAJOR COMPONENT).
 CC -|- TISSUE SPECIFICITY: NEWLY FORMED MESODERM AND LATER PROMINENTLY
 CC EXPRESSED IN HEMOCYTES, WHICH ALSO SYNTHESIZE COLLAGEN IV.
 CC -|- DEVELOPMENTAL STAGE: DURING MORPHOGENESIS, MOSTLY IN EMBRYO
 CC DEVELOPMENT AT 10-12 HOURS.
 CC -|- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -|- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
 CC -|- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -|- SIMILARITY: CONTAINS 21.5 LAMININ EGF-LIKE DOMAINS.
 CC -|- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV (DOMAIN IV) IS NOT
 CC SIMILAR TO LAMININ DOMAIN IV).
 CC -|- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M96388; AAA28662.1; -;
 DR EMBL; L07288; AAC37178.1; -;
 DR EMBL; M75882; AAA28661.1; -;
 DR HSSP; P02468; ITLE.
 DR FlyBase; FBgn0002526; LANA.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001886; LamNT.
 DR InterPro; IPR000034; Laminin_B.
 DR InterPro; IPR002049; Laminin_EGF.
 DR InterPro; IPR001791; Laminin_G.
 DR InterPro; IPR001230; Prenyltn.
 DR Pfam; PF00052; laminin_B; 1.
 DR Pfam; PF00053; laminin_EGF; 20.
 DR Pfam; PF00054; laminin_G; 5.
 DR Pfam; PF00055; laminin_Nterm; 1.
 DR PRINTS; PRO00011; EGFLAMIN.
 DR ProDom; PD002082; LamNT; 1.
 DR ProDom; PD003031; Laminin_B; 1.
 DR SMART; SM00180; EGF_Lam; 17.
 DR SMART; SM00001; EGF_Like; 1.
 DR SMART; SM00281; LamB; 1.
 DR SMART; SM00282; LamG; 5.
 DR SMART; SM00136; LamNT; 1.
 DR PROSITE; PS00022; EGF_1; 17.
 DR PROSITE; PS01186; EGF_2; 5.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF; 19.
 DR PROSITE; PS50025; LAM_G_DOMAIN; 5.
 DR Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 3712 LAMININ ALPHA CHAIN.
 FT DOMAIN 25 272 LAMININ N-TERMINAL (DOMAIN VI).
 FT DOMAIN 273 332 LAMININ EGF-LIKE 1.
 FT DOMAIN 333 402 LAMININ EGF-LIKE 2.
 FT DOMAIN 403 447 LAMININ EGF-LIKE 3.
 FT DOMAIN 448 494 LAMININ EGF-LIKE 4.
 FT DOMAIN 495 540 LAMININ EGF-LIKE 5.
 FT DOMAIN 541 586 LAMININ EGF-LIKE 6.

DOMAIN 587 631 LAMININ EGF-LIKE 7.
 DOMAIN 632 676 LAMININ EGF-LIKE 8.
 DOMAIN 677 731 LAMININ EGF-LIKE 9.
 DOMAIN 732 784 LAMININ EGF-LIKE 10.
 DOMAIN 785 815 LAMININ EGF-LIKE 11 (INCOMPLETE).
 DOMAIN 816 1374 DOMAIN IV'.
 DOMAIN 1375 1420 LAMININ EGF-LIKE 12.
 DOMAIN 1421 1465 LAMININ EGF-LIKE 13.
 DOMAIN 1466 1513 LAMININ EGF-LIKE 14.
 DOMAIN 1514 1564 LAMININ EGF-LIKE 15.
 DOMAIN 1565 1574 LAMININ EGF-LIKE 16 (N-TERMINAL).
 DOMAIN 1575 1775 LAMININ DOMAIN IV (DOMAIN IV).
 DOMAIN 1776 1808 LAMININ EGF-LIKE 16 (C-TERMINAL).
 DOMAIN 1809 1858 LAMININ EGF-LIKE 17.
 DOMAIN 1859 1916 LAMININ EGF-LIKE 18.
 DOMAIN 1917 1969 LAMININ EGF-LIKE 19.
 DOMAIN 1970 2016 LAMININ EGF-LIKE 20.
 DOMAIN 2017 2063 LAMININ EGF-LIKE 21.
 DOMAIN 2064 2111 LAMININ EGF-LIKE 22.
 DOMAIN 2112 2671 DOMAIN II AND I.
 DOMAIN 2672 2868 LAMININ G-LIKE 1.
 DOMAIN 2876 3048 LAMININ G-LIKE 2.
 DOMAIN 3055 3223 LAMININ G-LIKE 3.
 DOMAIN 3249 3528 LAMININ G-LIKE 4.
 DOMAIN 3534 3709 LAMININ G-LIKE 5.
 DOMAIN 2178 2249 COILED COIL (POTENTIAL).
 DOMAIN 2301 2321 COILED COIL (POTENTIAL).
 DOMAIN 2376 2450 COILED COIL (POTENTIAL).
 DOMAIN 2541 2676 COILED COIL (POTENTIAL).
 DOMAIN 3270 3296 POLY-THR.
 DISULFID 273 282 BY SIMILARITY.
 DISULFID 275 296 BY SIMILARITY.
 DISULFID 298 307 BY SIMILARITY.
 DISULFID 310 330 BY SIMILARITY.
 DISULFID 333 342 BY SIMILARITY.
 DISULFID 335 367 BY SIMILARITY.
 DISULFID 370 379 BY SIMILARITY.
 DISULFID 382 400 BY SIMILARITY.
 DISULFID 403 414 BY SIMILARITY.
 DISULFID 405 421 BY SIMILARITY.
 DISULFID 423 432 BY SIMILARITY.
 DISULFID 435 445 BY SIMILARITY.
 DISULFID 448 460 BY SIMILARITY.
 DISULFID 450 468 BY SIMILARITY.
 DISULFID 470 479 BY SIMILARITY.
 DISULFID 482 492 BY SIMILARITY.
 DISULFID 495 507 BY SIMILARITY.
 DISULFID 497 514 BY SIMILARITY.
 DISULFID 516 525 BY SIMILARITY.
 DISULFID 528 538 BY SIMILARITY.
 DISULFID 541 553 BY SIMILARITY.
 DISULFID 543 560 BY SIMILARITY.
 DISULFID 562 571 BY SIMILARITY.
 DISULFID 574 584 BY SIMILARITY.
 DISULFID 587 599 BY SIMILARITY.
 DISULFID 589 605 BY SIMILARITY.
 DISULFID 607 616 BY SIMILARITY.
 DISULFID 619 629 BY SIMILARITY.
 DISULFID 632 644 BY SIMILARITY.
 DISULFID 634 650 BY SIMILARITY.
 DISULFID 652 661 BY SIMILARITY.
 DISULFID 664 674 BY SIMILARITY.
 DISULFID 677 691 BY SIMILARITY.
 DISULFID 679 700 BY SIMILARITY.
 DISULFID 702 711 BY SIMILARITY.
 DISULFID 714 729 BY SIMILARITY.
 DISULFID 732 746 BY SIMILARITY.
 DISULFID 734 753 BY SIMILARITY.
 DISULFID 755 764 BY SIMILARITY.
 DISULFID 767 782 BY SIMILARITY.
 DISULFID 1375 1387 BY SIMILARITY.
 DISULFID 1377 1394 BY SIMILARITY.
 DISULFID 1396 1405 BY SIMILARITY.

FT	DISULFID	1408	1418	BY SIMILARITY.
FT	DISULFID	1421	1429	BY SIMILARITY.
FT	DISULFID	1423	1436	BY SIMILARITY.
FT	DISULFID	1438	1447	BY SIMILARITY.
FT	DISULFID	1450	1463	BY SIMILARITY.
FT	DISULFID	1466	1480	BY SIMILARITY.
FT	DISULFID	1468	1487	BY SIMILARITY.
FT	DISULFID	1489	1498	BY SIMILARITY.
FT	DISULFID	1501	1511	BY SIMILARITY.
FT	DISULFID	1514	1526	BY SIMILARITY.
FT	DISULFID	1516	1533	BY SIMILARITY.
FT	DISULFID	1535	1544	BY SIMILARITY.
FT	DISULFID	1547	1562	BY SIMILARITY.
FT	DISULFID	1859	1874	BY SIMILARITY.
FT	DISULFID	1861	1885	BY SIMILARITY.
FT	DISULFID	1887	1896	BY SIMILARITY.
FT	DISULFID	1899	1914	BY SIMILARITY.

Query Match 6.4%; Score 128.5; DB 1; Length 3712;
Best Local Similarity 23.4%; Pred. No. 24;
Matches 91; Conservative 48; Mismatches 155; Indels 95; Gaps 19;

Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL: Z46921; CAA87023.1; --
SGD: S0001431; YIL169C.
DR InterPro: IPR004089; Chemotaxis_transducer.
DR InterPro: IPR000727; T_SNARE.
KW Hypothetical protein; Signal.
CC
CC SIGNAL 1 23
FT CHAIN 24 995
FT TAG 24 995
FT CARBOHYD 28 28
FT CARBOHYD 35 35
FT CARBOHYD 35 35
FT CARBOHYD 468 468
FT CARBOHYD 664 664
SQ SEQUENCE 995 AA; 99735 MW; F63E287A03F137EC CRC64;

POTENTIAL.
HYPOTHETICAL PROTEIN YIL169C.
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 6.4%; Score 128; DB 1; Length 995;
Best Local Similarity 16.1%; Pred. No. 5.1;
Matches 73; Conservative 100; Mismatches 152; Indels 128; Gaps 15;

QY 4 SITKKNPNNGQTGLEARMELDTLINAAMTTLASLDQYAKIEASLSAYSEATVNNN 63
DB 136 SVSGSSASDVSSVSQSSASDVSSVSQSSASD-----VSSVSQSSASTSD 188
QY 64 LNAYLEQLMAKTNLESAINOANTDKTTFDNEHPLVEAYKALKTTLEQRAINLEGLS 123
DB 189 VSSVSQSSASDVSSVSQSSASDVSSVSQSSASTSDVSSVSQSSASTSGVSSS 248
QY 124 AYNQIRNLVDLYKASLITKTLDPLNGWLLDSNELTANKNNINNTLTINEQKTNAD 183
DB 249 GSQSVSS-----ASGSSSSFPQST-----SSASTASGATSNLSLSTSSASSAS 293
QY 184 ALSNSFIKK-----VIONEQSF---VG 203
DB 294 ATASNSUSSDGTIYLPITTTISGDLTLGKVIATGVVVVAAGAKLTLDGOKYSADLK 353
QY 204 TFTNANQPSNYSFVAFSADVTPVNYKARTVWNGDE---PSSRIILANTNSITDWSIY 260
DB 354 VYGDLLVKKSKETPGTEFDISGENFDVGN- FNAEASATSAISIYFTSPSSFDNSGDI 411
QY 261 SLA-GTNTKYQFSPNSGPGTGYLYFPYKLVKAADANNVGLQYKLNNGVQVQVFAST- 318
DB 412 SLSLSKSKKGTEVTFPSYNSGAFS-----SNAI-----LNGSVSGLQRDDTE 456
QY 319 -SANTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYL- 374
DB 457 GSVNNGEIN-----LDNGSTYVIVPEVPVGKGTVNIIS---GNLYLHP 496
QY 375 -----SSNENNADKIP--GY 387
DB 497 DTFTCQTVVFKGGLAVDPTETNATPIPVVGY 529

RESULT 12
YEF3_YEAST
ID YEF3_YEAST STANDARD; PRT; 956 AA.
AC P32618;
DT 01-OCT-1993 (Rel. 27, Created)
DR 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 106.1 kDa protein in GLY1-GDA1 intergenic region.
GN YEL043W OR SVGP-ORF14.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

QY 278 -----PSTGYLYFPYKLVKAADANNVGLQYKLNNGNVQ 310
 DB 678 KGFTTDELNDYTKQOPVRSSTNESLFTSGTPMSSYK-----ANPVISPY--SSHLR 729
 QY 311 QVEFATSTANNTTAMPTPAVDIKVAKIVLSGLRF-----GQNTIELS-- 354
 DB 730 QTSNATNTNPMHPQSLAATLNDPSLQSFVRSGFSYSAFPANSLQNNINGNETENISPR 789
 QY 355 -----VPTGEGNNKVPAMIGNIYLSSNENNADKI 384
 DB 790 ISSDFNLLVNLSPRLSNVDPIVPGNNTLTTPSHSNLLTNHQTADNI 838

RESULT 13
 PSTL_YEAST STANDARD; PRT; 444 AA.
 AC Q12355;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Protoplast secreted protein 1 precursor.
 GN PST1 OR YDR055W OR D4214 OR YD9609.09
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 NCBI_TaxID=4932;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96381250; PubMed=8789263;
 RA Brandt P., Ramlow S., Otto B., Bloeker H.;
 RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm
 of Saccharomyces cerevisiae chromosome IV.";
 RL Yeast 12:85-90(1996).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;
 RL Submitted (May 1995) to the EMBL/GenBank/DBJ databases.
 [3]
 RN IDENTIFICATION.
 RP MEDLINE=99251092; PubMed=10234784;
 RA Pardo M., Monteoliva L., Pla J., Sanchez M., Gil C., Nombela C.;
 RT "Two-dimensional analysis of proteins secreted by Saccharomyces
 cerevisiae regenerating protoplasts: a novel approach to study the
 cell wall.";
 RL Yeast 15:459-472(1999).
 [4]
 RN GPI-ANCHOR.
 RP MEDLINE=20469049; PubMed=11016834;
 RA Terashima H., Yabuki N., Arisawa M., Hamada K., Kitada K.;
 RT "Up-regulation of genes encoding glycosylphosphatidylinositol
 (GPI)-attached proteins in response to cell wall damage caused by
 disruption of FKS1 in Saccharomyces cerevisiae.";
 RL Mol. Gen. Genet. 264:64-74(2000).
 CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR AND
 SECRETED BY REGENERATING PROTOPLASTS.
 CC -!- SIMILARITY: BELONGS TO THE SPS2 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; X84162; CAA58971.1;
 EMBL; Z74351; CAA98873.1;
 EMBL; Z49209; CAA89084.1;
 SGD; S0002462; PST1.
 COMPLEYEST-2DPAGE; Q12355;
 Glycoprotein; Membrane; GPI-anchor; Signal.
 SIGNAL 1 19
 POTENTIAL.

OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,
 Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
 Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
 Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 [3]
 RN FIBRONECTIN TYPE III DOMAIN.
 RP MEDLINE=97148176; PubMed=8994808;
 RA Bateman A., Chothia C.;
 RT "Fibronectin type III domains in yeast detected by a hidden Markov
 model.";
 RL Curr. Biol. 6:1544-1546(1996).
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; U18779; AAB64999.1;
 DR PIR; S30834; S30834.
 DR SGD; S0000769; YEL043W.
 DR InterPro; IPR003961; FN_III.
 DR Pfam; PF00041; fn3; 1.
 DR SMART; SM00060; FN3; 1.
 KW Hypothetical protein.
 FT DOMAIN 35 125 FIBRONECTIN TYPE-III.
 SQ SEQUENCE 956 AA; 106132 MW; 3F78B09A0FCA03AF CRC64;

Query Match 6.3%; Score 127; DB 1; Length 956;
 Best Local Similarity 19.7%; Pred. No. 5.5;
 Matches 104; Conservative 86; Mismatches 163; Indels 176; Gaps 25;

QY 5 ITRKDANPNNGQTOLEAARMELTDL-----INA-RMTLASLDQYAKIEASLSAYSEA 57
 DB 337 LNESASVANINKIEISLONEISKEESNRKLNASKSLITSIVNANVENDKPIASGEL 396
 QY 58 ETYNNLN-ATLE-----QLKMAKTLESAINANTDKTTFDNEHNLVEAYKALK 107
 DB 397 SAVLKLNKLDFTLENGFSLNAGEEFLSKLNADSLIKWIKQELSIDQE---LEANWKLQR 453
 QY 108 TLEQRATNLEGLSTAYNOIRNLVDLYNKASSLIPTKL-DPL--NGHTLDSNETTTA 164
 DB 454 SNLLKKISALE-----NQF--NEMSLNLR--NLKTKLMVQPYKKNNGDSLAATNSNSA 502
 QY 165 KKN-----INNTLSTINEQKTNADAL----- 185
 DB 503 EKNRSSGSIQLPLSNMSTRGTSIDLSNNKSN--NSNADSAAPLRLHNPVSPSPNEP 560
 QY 186 ---SNSETKKIQNNE--OSFVGTETNAN-----VQPSNYSFVA-----FSADYTPVNYK 230
 DB 561 IQPSSLSQLTQTDIDNLSMLSNHSSNNENKQPPSSYSHALPTATANATATATATN-G 619
 QY 231 YARRTVWNG---DEPSSRLA-----NTNSITDVSIMYSLAGTNTKYQSFNSYNG 277
 DB 620 HSRSLNLTWTAQAPQSHQOVSTELDQAFEDYDNANHL--ISGLQNMIDYDTPDNISNYS 677

FT	CHAIN	20	?	PROTOPLAST SECRETED PROTEIN 1.
FT	PROPEP	?	444	REMOVED IN MATURE FORM (POTENTIAL).
FT	DOMAIN	356	416	SER-RICH.
FT	CARBOHYD	57	57	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	76	76	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	83	83	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	86	86	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	196	196	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	210	210	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	228	228	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	235	235	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	242	242	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	263	263	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	268	268	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	280	280	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	292	292	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	305	305	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	329	329	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SEQUENCE	444 AA;	45776 MW;	230F60CACA5921A4 CRC64;

Query Match	6.3%;	Score 126.5;	DB 1;	Length 444;
Best Local Similarity	18.1%;	Pred. No. 2.3;		
Matches	73;	Conservative	79;	Mismatches 146; Indels 105; Gaps 15;

QY	17	TQLEAARWE-LTDLINAKAMTSLASLDYAKTEASLSAYSEAEVNNNNLNATLQOLKMAK	75
DB	88	TSPAADLESITDSLNQSLTILT-----SASFGSLQVSDSIKLTLPALISSFT	136
QY	76	TNLESAINQANTOKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSS--TAYNOIRNLV	133
DB	137	SNIKSANIIVISDTSL-----QSVDFGSALKVYVFNVNNKKLTSLKSPVEIVSDQLQ	190
QY	134	DLYNKASSLITKTLDPPL---NGGTLDSNEITTANKN-----INNTLSTINEQKINA	182
DB	191	FSFNGNQTKI--TFDDLWANNISLTDVHVSVFANLQKINSSLSGLFINNSSLNFTKLNT	248
QY	183	-----DALSNSFIKK-----VTQNN-----BQSPVGFPTN	207
DB	249	IGOTFTSVINDYLNLSFNLSLSTIGGALVANNNTGLQIGGLDLNLTIGGTLEVVGNFTS	308
QY	208	ANYQPSNYSVAFSADYTPVNYKYARRTVWNGDEPSSRLANTNSITDVSWIYSLAGTNT	267
DB	309	LNL--DSLKSVKGADVESKSNFSCHAL-----KALQKGGIKGEVFCVNGASST	358
QY	268	KYQFSFNGYSPGTGLYFPYKVKAAADANNVGLQKLNNGVQVQVEFATSTSANNTANP	327
DB	359	SVKLSSTSKSQSS-----QTTAKVSKSSKAEKKFTSGDIKAAASASSVSSSGASS--	410
QY	328	TPAVDEIKVAKIVLGLURFGONTIELSVPTGEGNNKNVAPMIG	370
DB	411	-----SSSKSSKGNAIIMAPIGO-----TTPLVG	434

RESULT	14	
MSPI_1_PLAIFK		
ID	MSPI_1_PLAIFK	STANDARD; PRT: 1630 AA.
AC	P04932;	
DT	13-AUG-1987 (Rel. 05, Created)	
DT	01-FEB-1996 (Rel. 33, Last sequence update)	
DT	01-OCT-1996 (Rel. 34, Last annotation update)	
DE	Merozoite surface protein 1 precursor (Merozoite surface antigens) (PMM5A) (P190).	
GN	MSP-1.	
OS	Plasmodium falciparum (isolate K1 / Thailand).	
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
NCBI_TaxID	5839;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE-86135024; PubMed-3004972;	
RA	MacKay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,	
RT	Stunnenberg H., Bujaard H.;	
RT	*polymorphism of the precursor for the major surface antigens of	

Plasmodium falciparum merozoites: studies at the genetic level.";
EMBO J. 4:3823-3829(1985).
[2]
REVISIONS, SEQUENCE FROM N.A.
RN Pen W., Tolle R., Bujard H.;
RP Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RL
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (potential).
CC
CC -1- PTM: MEROZOITE SURFACE ANTIGEN CONTAIN THE SEQUENCE OF 83 kDa, 42
CC kDa and 19 kDa ANTIGENS WHICH ARE THE MAJOR SURFACE ANTIGENS OF
CC MEROZOITES. THE MATURATION TAKE PLACE DURING SCHIZONT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X03371; CAA27070.1; -.
DR PIR; A25120; SAZQK1.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor.
FT SIGNAL 1 19 POTENTIAL
FT CHAIN 20 1630 MEROZOITE SURFACE PROTEIN 1.
FT DOMAIN 67 84 TRIPEPTIDE SG(TP) REPEAT.
FT TRANSMEM 1614 1630 MEMBRANE ANCHOR.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 755 755 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 759 759 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 774 774 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 835 835 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 955 955 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1630 AA; 187289 MW; ADBDEC3CE0A46322 CRC64;

Query Match 6.3%; Score 126.5; DB 1; Length 1630;
Best Local Similarity 20.7%; Pred. No. 11;
Matches 86; Conservative 62; Mismatches 155; Indels 113; Gaps 19;

Qy	37	LASLDYAK-----	IRASUSSAYSEAE	TWNNNI	NATLEQ	LKMAK	TWLESA	INQANTDKTT	91
		: : :	:	:	:	:	:	:	
Db	226	VGMEDYIKKNKKT	INENIELIEESKKT	IDKKN	NATKEEK-----	KKLIQAQYDL	SL	278	
Qy	92	FDNEHPNLV	EAYKAL	TTLEORAT	NLEGUSSTAY	NOIRNN	LDVLYNK	KASLITK	LDPLN
		: : :	:	:	:	:	:	:	
Db	279	YNKQ---LEE	AHN-LISV	EKR	DTLTK-----	NENIKEL	LKINEI--	KNPPAN	323
Qy	152	GG---TL	DSN-EITT	ANKNIN	NTLSTINE	OQT	NADAL-----	SNSFTK	191
		:	:	:	:	:	:	:	
Db	324	SGNTPN	LLDKNK	KIEBEKE	IEIAKTI---	KFNIDS	LFTD	PLEY	YLREKNNK
		:	:	:	:	:	:	:	
Qy	192	KVQNN	EQSGVGT	TIN	ANOVPSY	SFV-----	AFSAD	VPVNYKY	ARRTWNGDEPS
		: : :	:	:	:	:	:	:	
Db	381	AKVET	KESTPE	NYPGV	TYPLSY	DINNAL	NELNSFGD	LIN	PDFYT-----
		: : :	:	:	:	:	:	:	
Qy	244	SRILAN-----	TNSIT	DVSN	WISY	SLAGNT	WIKYQS	272	
		:	:	:	:	:	:	:	
Db	432	KNIVTD	NERK	FFINE	IKIEKK	KIESDK	SVEDRSK	SLNDIT	KEYEKL-LNEIY
		:	:	:	:	:	:	:	
Qy	273	FSNYG	STGY-----	LYFP	PKL	KAADAN	NVGLQY	KLINGN	WQVFEATST
		:	:	:	:	:	:	:	
Db	491	FNN	IDLTPN	EKM	GKRRYSY	KEV	KLTH-HNT	FASYEN	SKHNLEK
		:	:	:	:	:	:	:	

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 12, 2002, 10:50:40 ; Search time 48.58 seconds
(without alignments)
1403.047 Million cell updates/sec

Title: US-09-147-052-4_COPY_693_1086

Perfect score: 2004

Sequence: 1 GCMSITKDKANPNNGQTGLE.....SSNNADKIPGYRRPGTGL 394

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1200	59.9	271	2 Q49464	Q49464 mycoplasma
2	837	41.8	671	2 Q9L8D6	Q9L8D6 mycoplasma
3	831.5	41.5	702	2 Q49499	Q49499 mycoplasma
4	826	41.2	632	2 Q9XCG8	Q9XCG8 mycoplasma
5	805	40.2	584	2 Q9KH13	Q9KH13 mycoplasma
6	761	38.0	680	2 Q9KH14	Q9KH14 mycoplasma
7	754	37.6	702	2 Q49498	Q49498 mycoplasma
8	746	37.2	645	2 Q9ZID1	Q9ZID1 mycoplasma
9	745	37.2	644	2 Q9ZHR9	Q9ZHR9 mycoplasma
10	744	37.1	650	2 Q49495	Q49495 mycoplasma
11	743	37.1	649	2 Q49497	Q49497 mycoplasma
12	713	35.6	647	2 Q49468	Q49468 mycoplasma
13	704	35.1	656	2 Q9KH15	Q9KH15 mycoplasma
14	690.5	34.5	703	2 Q05122	Q05122 mycoplasma
15	619.5	30.9	419	2 Q9L8D5	Q9L8D5 mycoplasma
16	575.5	28.7	386	2 Q49500	Q49500 mycoplasma

17	557	27.8	367	2 Q9XCG7	Q9XCG7 mycoplasma
18	189.5	9.5	320	2 Q49496	Q49496 mycoplasma
19	175	8.7	6713	16 Q99054	Q99054 staphylococ
20	175	8.7	6713	16 Q931R6	Q931R6 staphylococ
21	173	8.6	1302	2 Q49547	Q49547 mycoplasma
22	164.5	8.2	2481	16 Q99QR6	Q99QR6 staphylococ
23	163.5	8.0	3890	16 Q99U53	Q99U53 staphylococ
24	160.5	8.0	433	16 Q97TP6	Q97TP6 clostridium
25	160.5	8.0	682	2 Q48674	Q48674 lactococcus
26	159	7.9	933	2 Q53653	Q53653 staphylococ
27	158.5	7.9	1344	2 Q49545	Q49545 mycoplasma
28	156	7.8	807	5 Q96262	Q96262 plasmodium
29	155	7.7	1072	16 Q9CF64	Q9CF64 lactococcus
30	155	7.7	4688	16 Q9PQ08	Q9PQ08 ureaplasma
31	154.5	7.7	1365	2 Q49525	Q49525 mycoplasma
32	153.5	7.7	1795	2 Q9LJC9	Q9LJC9 staphylococ
33	153.5	7.7	2478	2 Q9RL69	Q9RL69 staphylococ
34	151.5	7.6	1237	16 Q9ZKS7	Q9ZKS7 helicobacte
35	151.5	7.6	2478	2 Q9LCH2	Q9LCH2 staphylococ
36	151	7.5	624	2 Q49548	Q49548 mycoplasma
37	150.5	7.5	604	5 Q26021	Q26021 plasmodium
38	150.5	7.5	1051	2 Q49524	Q49524 mycoplasma
39	149.5	7.5	962	2 Q49546	Q49546 mycoplasma
40	148	7.4	716	2 Q49526	Q49526 mycoplasma
41	147.5	7.4	655	5 Q26109	Q26109 plasmodium
42	146	7.3	654	2 Q9RLP9	Q9RLP9 mycoplasma
43	145	7.2	2399	16 Q9ZKS9	Q9ZKS9 helicobacte
44	144	7.2	1107	2 Q9F2D8	Q9F2D8 salmonella
45	143.5	7.2	751	3 Q94358	Q94358 schizosacch

ALIGNMENTS

RESULT 1

Q49464	PRELIMINARY;	PRT;	271 AA.
ID Q49464;			
AC Q49464;			
DT 01-NOV-1996 (TRENBLrel. 01, Created)			
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)			
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)			
DE TM-1 (FRAGMENT).			
GN TM-1.			
OS Mycoplasma gallisepticum.			
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;			
OC Mycoplasmataceae; Mycoplasma.			
OX NCBI_TaxID=2096;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=94025893; PubMed=8212828;			
RA Saito S., Fujisawa A., Ohkawa S., Nishimura N., Abe T., Kodama K.,			
RA Kanogawa K., Aoyama S., Iritani Y., Hayashi Y.;			
RT "Cloning and DNA sequence of a 29 kilodalton polypeptide gene of			
RL Mycoplasma gallisepticum as a possible protective antigen.";			
RL Vaccine 11:1061-1066(1993).			
DR EMBL: S65869; AAB28343.2; -			
FT NON_TER	271	271	
SQ SEQUENCE	271 AA;	29817 MW;	8B25DE0CD5C85CA2 CRC64;

Query Match 59.9%; Score 1200; DB 2; Length 271;
Best Local Similarity 97.1%; Pred. No. 2.2e-51;
Matches 238; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy	2	CMSITKDKANPNNGQTGLEAARMELTDLINAKAMTSLASLDYAKIEASLSAYSEATVN	61
Db	27	CMSITKDKANPNNGQTQLQAARMELTDLINAKARTLASLDYAKIEASLSAYSEATVN	86
Qy	62	NNLNATLEQLKMAKNLNSAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS	121
Db	87	NNLNATLEQLKMAKNLNSAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA	146
Qy	122	STAYNQIRNNLVLDLYNKASSLITKTLDPLNGCTLLDSNEITTANKNINTLTSTINEQKTN	181

||||| 147 STAYNOIRNNLDVLYNNASSLITLPLNGGMLDSNEITVRNINNLSTINEOKTN 206
QY 182 ADALSNFIFKVIQNNQSGVGTFTNANQPSNYSFVAFSADVTVPVNYKYARRTVWNGDE 241
Db 207 ADALSNFIFKVIQNNQSGVGTFTNANQPSNYSFVAFSADVTVPVNYKYARRTVWNGDE 266
QY 242 PSSRI 246
Db 267 PSSRI 271
RESULT 2
Q9L8D6 PRELIMINARY; PRT; 671 AA.
AC Q9L8D6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PMGA-LIKE PROTEIN 9.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F;
RA Pharr G.T., Branton S.L., Hanson L.A., Minion F.C., Lott B.D.,
RA May J.D., Hughlett M.B.;
RT "A novel PMGA-like gene from the F-strain (vaccine strain) of
RT Mycoplasma gallisepticum";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF210770; AAF29524.1; -;
SQ SEQUENCE 671 AA; 71898 MW; 4D6AE6B59175D679 CRC64;
Query Match 41.8%; Score 837; DB 2; Length 671;
Best Local Similarity 43.4%; Pred. No. 2.5e-33;
Matches 179; Conservative 75; Mismatches 118; Indels 40; Gaps 9;
QY 9 DANP-----NNGQTLEARMELTDLINAKAMTASLDQYAKIEASLSAYSEAEVTNN 62
Db 55 DTNPGNDGGMENSAQOAAAKKELSDLLATQSNLAKAYADYTNQNTLTAAVTTAKSTD 114
QY 63 NLNATLEQKMAKTNLESAINOANTDKTTFDNEHPNLVEAYKALKTTLRQATNLEGLSS 122
Db 115 NTSVLEQVKSATSTLQAADTAASKTSFDEKPELIRKAYALKETLKNETVLSGLTD 174
QY 123 TAYNOIRNNLDVLYNKASSLITKTLDPLNGGTLDSNEITTANKNNINLTSTINEOKTNA 182
Db 175 SNFATIKTNLTALYQSGKDFVKATLDPVSGNA-PQIADITKADKIDADAVSKLETWKTNA 233
QY 183 DALSNFIFKVIQNNQSGVGTFTNANQPSNYSFVAFSADVTVP-----VNYKYARRTVN 238
Db 234 NTLATSFVKELVKNNLTGIDT-TNNRQPGNYSFVGYSVNATNNNEINFWNAQRKVT 292
QY 239 GDEPSSRIIA----NTNSTIDVSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVKAAD 294
Db 293 SDNGRTLSISSDNDSTITEVSWIYSLGAGTKYSLITFNYYGPSTGYLYFPYKLVKEGD 352
QY 295 ANNVLQYKLNGNNGVQVEFA-----TSTSA-----NNTT-----ANPTPAVD 332
Db 353 ENNVALQYTLNLSGSAQEVNFAPTVKTSVSADSGSDSNNOTESAAETMPVTSDLNPAPTWS 412
QY 333 EIKVAKIVLSGLRFGONTIELSVPTGEGNNKVPAMIGNIYLSNNENNADKI 384
Db 413 DINIAKLTLSNLKFGSNTIEFSVPTPEPS--NKVAPMIGNMYLTSTIANEAKV 462
RESULT 3
Q49499 PRELIMINARY; PRT; 702 AA.
ID Q49499
AC Q49499;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PMGAI.4 PROTEIN PRECURSOR.
GN PMGAI.4.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "the organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum";
RL FEBS Lett. 352:347-352(1994).
DR EMBL; L28424; AAA62418.1; -;
KW SIGNAL.
FT SIGNAL 1 25 POTENTIAL.
SQ SEQUENCE 702 AA; 75517 MW; B70AC874FE85055C CRC64;
Query Match 41.5%; Score 831.5; DB 2; Length 702;
Best Local Similarity 42.0%; Pred. No. 4.8e-33;
Matches 173; Conservative 84; Mismatches 116; Indels 39; Gaps 8;
QY 9 DANPNG-----QTQLEARMELTDLINAKAMTASLDQYAKIEASLSAYSEAEVTNN 62
Db 68 NTPNGNGGGTDNAQOAAAKKELSDLLATQSNLSTYADYANIONTLTAATYTTAKSTD 127
QY 63 NLNATLEQKMAKTNLESAINOANTDKTTFDNEHPNLVEAYKALKTTLRQATNLEGLSS 122
Db 128 NTSVLEQVKSATSTLQAADTAASKTSFDEKPELIRKAYALKETLKNRNSLSGLTD 187
QY 123 TAYNOIRNNLDVLYNKASSLITKTLDPLNGGTLDSNEITTANKNNINLTSTINEOKTNA 182
Db 188 SNFATIKTNLTALYQSGKDIVTKTLDPLM-GTAINLSAYSQANTNISNAVSKLETWKTNA 246
QY 183 DALSNFIFKVIQNNQSGVGTFTNANQPSNYSFVAFSADVTVP-----VNYKYARRTV 236
Db 247 TVLATSFVKELVKNNLTGIDT-TNNRQPGNYSFVGYSDVTGSDNAPNWSFAQRKV 305
QY 237 WNGD-----EPSRILANTNSITDYSWIYSLAGTNTKYQSFNSYGPSTGYLYFPYKLVK 291
Db 306 WTSNTDILSQPOPAEGENQOASAPDVSWIYNLTGMAKAKYSLITFNYYGPSTGYLYFPYKLVN 365
QY 292 ADANNVGLQYKLNGNNGVQVEFATS-----TSANN-----TTANPTPAVD 332
Db 366 SSDSDKVALLEYKLNSAVKTIDFSPQTSFVADATRENNRSTAAAPAQGSTENPAPTLD 425
QY 333 EIKVAKIVLSGLRFGONTIELSVPTGEGNNKVPAMIGNIYLSNNENNADK 383
Db 426 DIKIAKLTLSNLKFGSNTIEFSVPTTAKETSKVAPMIGNMYLTSSDROVYNK 477
RESULT 4
Q9XCG8 PRELIMINARY; PRT; 632 AA.
AC Q9XCG8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE VLHAI PRECURSOR (FRAGMENT).
GN VLHAI.
OS Mycoplasma imitans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=29560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4229;

```

RX MEDLINE=99392472; PubMed=10463176;
RA Markham P.F., Duffy M.F., Giew M.D., Browning G.F.;
RT "A gene family in Mycoplasma imitans closely related to the pmga
FT family of Mycoplasma gallisepticum.";
RL Microbiology 145:2095-2103(1999).
DR EMBL; AF141940; AAD39483.1; -.
KW Signal.
FT NON_TER 1
FT SIGNAL <1 17 POTENTIAL.
FT CHAIN 18 632 VLHAI
FT SEQUENCE 632 AA; 66959 MW; 1735B12E705BE47 CRC64;

Query Match 41.2%; Score 826; DB 2; Length 632;
Best Local Similarity 45.3%; Pred. No. 8e-33;
Matches 183; Conservative 65; Mismatches 120; Indels 36; Gaps 9;

Qy 8 KDAPNNQGT-----OLEAARMELTDLINAKAMTILASLDQYAKIEASLSAYSAE 58
Db 34 KPNPDTGNTNPGDGTDAKQADAKASLNTLLGSQSTNVALYEDYAKIKDTLSSAYSAQ 93

Qy 59 TVNNLNATLEQKMAKTNLESATNQANTDKTTFDNEHPNLVEAYKALKTTLEQRTATNLE 118
Db 94 TTANNANATLKDNDAAKALQTAISDAANAKTEFDKANGLVSAAYAKUKETLKSETTNDL 153

Qy 119 GLSSTAYNQIRNNLVLYNKASSLITTKTLDPLNGGTLSDSNEITTANKNNINNTLSTINEQ 178
Db 154 GLSQDYSIAKANLVSLYKAKDKTTVTLDLP-TSGMIPKVDIEITSANTAIQAYSAIDSQ 212

Qy 179 KTNADALSNSFIKKVQIONNEQSFVGTFTN-----ANVQPSNYSFVAFSADVTP----- 226
Db 213 KTNADTAATTEIKELKLSAKLT-PGTTAEGQQAASQPGNYSFVGFSDVTTGRTGSQEQ 271

Qy 227 --VNYKYARRVTWNGDEPS---SRLANTNSITDVSWIYSLAGTNTKYQFSFNGYGPSTG 281
Db 272 DLPSWNFAKRVWTEGLSAGQTLVSSSTPLTLDVSWIYSLTGAGSKYTLTFTTYGPTA 331

Qy 282 YLYFPYKLVKAADANNVGLQKLNNG-NVQOVEFATSTSANNTTANPTPAYDEIKVAKIV 340
Db 332 YLYFPYKLVQSSDSNKIGLOVKLNNSELTLPITFGNETN-----DSGATPADDIKVEKVT 387

Qy 341 LSGLRFQNTGTELSVPTGEGNNKVPAMIGNIYSSNENNAKDI 384
Db 388 LSNLNFQDNTIEFSVATEE---NKVAPMIGNNYLTSTSTNNVDKI 428

RESULT 5
Q9KH13 PRELIMINARY; PRT; 584 AA.
AC Q9KH13;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE ADHESIN PMGAL.4 (FRAGMENT).
GN PMGAL.4.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmga multigene family of Mycoplasma
RL gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275312; AAF91415.1; -.
DR InterPro; IPR001986; BPSP_syntase.
DR PROSITE; PS00104; BPSP_SYNTASE_1; UNKNOWN_1.
FT NON_TER 584
FT SEQUENCE 584 AA; 62453 MW; 5C467BA55FB27A72 CRC64;

```

```

Query Match 40.2%; Score 805; DB 2; Length 584;
Best Local Similarity 42.4%; Pred. No. 7.6e-32;
Matches 180; Conservative 72; Mismatches 113; Indels 60; Gaps 10;

Qy 11 NPNNQGT-----OLEAARMELTDLINAKAMTILASLDQYAKIEASLSAYSEATVNNLNA 66
Db 69 NPSNGTTPTEQQLAAARKTLTDLTGENTNVALYADYAKIQSTLSTAYMTAKTASENTSA 128

Qy 67 TLEQLKMAKTNLESATNQANTDKTTFDNEHPNLVEAYKALKTTLEORATNLEGLSSTAYN 126
Db 129 TLDNLSASTTLQAALDRAASNKRTFDSANQPLVTAYNQLKTTLOSKTTSLEGLSENKY 188

Qy 127 QIRNNLVLYNKASSLITTKTLDPLNGGTLSDSNEITTANKNNINNTLSTNEQKTNADALS 186
Db 189 SIKNHLKSLFDAGSAIAARTLDP-TMGIVPEVMSYTKANEDIMTAVSKLTEKTNADKF- 246

Qy 187 NSFIKKVTQNNQSFVGTFTNAN--VQPSNYSFVAFSADVTP-----VNYKYARTT 235
Db 247 NDFEKKPL--SKEKLVSTNDRAHNOEQPANWFSAGYSVDLTGTGSGNSQNLPNWNFQARK 304

Qy 236 VYNGD---EPSSRIANTNSITDVSWIYSLAGTNTKYQFSFNGYGPSTGYLYFPYKLVKA 292
Db 305 VTSSEQQTGKTALVSSPVSATDVSWIYSLAGETKYLTSFEYGPDTAFLIFPKLVKQ 364

Qy 293 ADANNVGLQYKLNNGNVQVEP-----ATST-----S 319
Db 365 ADSSSVALQYSLNKTSSKLINEPAKTPMTNADQSENGVATTSTTEGRSSSEVLVADEVA 424

Qy 320 ANNTTANPTPAYDEIKVAKIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYSSNEN 379
Db 425 AVNEMNPTPTVDINIAKVTLSGLTFGENTIEFSVPT-----NKVAPMIGNNYLTSSNG 479

Qy 380 NADKI 384
Db 480 SQGI 484

RESULT 6
Q9KH14 PRELIMINARY; PRT; 680 AA.
AC Q9KH14;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE ADHESIN PMGAL.3.
GN PMGAL.3.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmga multigene family of Mycoplasma
RL gallisepticum strain HS.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275312; AAF91414.1; -.
SQ SEQUENCE 680 AA; 72909 MW; A2ABC6500E3DFE47 CRC64;

Query Match 38.0%; Score 761; DB 2; Length 680;
Best Local Similarity 39.8%; Pred. No. 1.2e-29;
Matches 164; Conservative 71; Mismatches 129; Indels 48; Gaps 7;

Qy 9 DANPNNGQTQLEAARME-----LTDLINAKAMTILASLDQYAKIEASLSAYSEATVYN 61
Db 63 DTPNGGGQGMNATNQELVNAKKAALSDLLGGESKVELYADYAKIKADITSAAYAKTTS 122

Qy 62 NNLNATLEQKMAKTNLESATNQANTDKTTFDNEHPNLVEAYKALKTTLEQRTATNLEGLS 121
Db 123 DSSSTLDQVKTATSTLQTAINTAASDKKFDQSQSQLLMKAYKVLKDLINKKEAIVMSLN 182

```

QY 122 STAYNQIRNNLVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNNNTL-----ST 174
DB 183 QEKYAILSEINAASTAEIEVQKTLNVPNG-----NLPVVAALNAETKILEAIKEEK 236
QY 175 INEQKTNADALSNFPIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPVNYKYARR 234
DB 237 INSEKSNADLAFANYQYKL---DRTKLMSEGSNNTKQPCNYSFVAYASDIASPNWFAQR 293
QY 235 TVWNGD-----EPSSRIILANTNSITDVSNIWISLAGTNTKYOFSPNSYGPSTGYLYFPYKL 289
DB 294 TVWTADSRWTSPPLNNLQNSAPLTDVSWITLSGTGAKYTLTDFYDYGPTQGYLYFPYKL 353
QY 290 VKAADANNVGLQYKLNNGVQVEF-----ATSTANNTTANPTAV 331
DB 354 VKTSD--KVGLQYKLNQADPVAIQSEAAATASAPAEATDGRQESAEATATANEKVNMPSPV 411
QY 332 DRIKVAKTVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYLSNNENADK 383
DB 412 NTINAKVTVLSNLKFGSNTIEFSPMDQDNMKVPAMIGNIYITSNDEANK 463
RESULT 7
Q49498
ID Q49498 PRELIMINARY; PRT: 702 AA.
AC Q49498:
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE PMGAL.3 PROTEIN PRECURSOR.
GN PMGAL.3
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Giew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum.";
RL FEBS Lett. 352:347-352(1994).
DR EMBL: L28424; AAA62417.1; -.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 702 AA; 75537 MW; 273E8915FE57B9F CRC64;

Query Match 37.6%; Score 754; DB 2; Length 702;
Best Local Similarity 41.6%; Pred. No. 2.8e-29;
Matches 176; Conservative 62; Mismatches 127; Indels 58; Gaps 11;
QY 11 NPNNGOT----QLEARMELTDLINAKMTLASQDYAKIEASLSAYSEAEIVNNLNA 66
DB 69 NPNPGTTPPEQLAARAKTLTDLGENTNVALYADYAKIQSTLSYATWAKTASNTSA 128
QY 67 TLEQLKMAKTNLESAINQANTOKTTFDNEHPNLVEAYKALKTTLEORATNLEGLS 126
DB 129 TLENLSASTTLQAAIDKAANDKRVFDSYQNPVLAAYNNLKTTLTKSKTTSLEGLSEN 188
QY 127 QIRNNLVLYNKASSLITKTLDPLNG--GTLLDSNEITTANKNNNTLS--TINEOKTNA 182
DB 189 GIKNHLKSLFDGSAITAKTLDPTSGERTLEKVN---ANNCKIAWISPELSKKWKGNA 245
QY 183 DALNSFIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTP-----VNYKYARTVW 237
DB 246 DKF-NEFEKNPLSKELKSTSDTAHQEQPAPNWSFAAAYSDLTNSQNLPLNWNFAORKVW 304
QY 238 NGD--EPSSRIILANTN--SITDVSNIWISLAGTNTKYOFSPNSYGPSTGYLYFPYKL 294
DB 305 TSENQOFGKALVSSPVSATFDVSNISWISLAGEGTKYTLTTEYIGPDNFAFLYLPYKLK 364

QY 295 ANNVGLQYKLNNGVQVEF-----ATST-----SAN 321
DB 365 SSSVALQYSLNKTSSKLNFLNPAETVSTNTQSENEVATTSTEARSSYKVLVDAEAT 424
QY 322 NTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYLS 381
DB 425 NNMNHTTVDSDINIATKVTLSGLTGTGENTIEFSVPEG-----KVAPMIGNMYLTS 479
QY 382 DK1 384
DB 480 VK1 482
RESULT 8
Q9ZIDI
ID Q9ZIDI PRELIMINARY; PRT: 645 AA.
AC Q9ZIDI:
DT 01-MAY-1999 (TremBLrel. 10, Created)
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TremBLrel. 10, Last annotation update)
DE M9 PROTEIN.
GN M9.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31; ATCC19610;
RX MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (M9) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmga
RT family.";
RL Infect. Immun. 66:5570-5575(1998).
DR EMBL: AF032890; AAC69269.1; -.
SQ SEQUENCE 645 AA; 69876 MW; 5A2F29572CFAB055 CRC64;

Query Match 37.2%; Score 746; DB 2; Length 645;
Best Local Similarity 42.3%; Pred. No. 6.2e-29;
Matches 172; Conservative 64; Mismatches 123; Indels 48; Gaps 11;
QY 9 DANPNNGQ-----TQLEARMELTDLINAKMTLASQDYAKIEASLSAYSEAEIVN 61
DB 44 DTNPGDGGGGMNNAASOELAAARMGLTTFVDSRAKNLGLVVDYKTKQDTLTAKYDAKT 103
QY 62 NNLNATLEOLKMAKTNLESAINQANTOKTTFDNEHPNLVEAYKALKTTLEORATNLEGLS 121
DB 104 DNSSTTQNLNEAKTRLETAIRTAATSKQTFDEQHAELVKVYEELKTTLSNETATLAPYA 163
QY 122 STAYNQIRNNLVLYNKASSLITKTLDPLNGGTLSDSNEITTANKNNNTL--STINEOK 179
DB 164 AAQYAGIKMHLISGLYDAGKAITTKTLEPVEGDP--LPADVMMANTKIVEAIKDEVLP 222
QY 180 TNADALSNIKKVIONNEQSFVGTFTNANVQPSNYSFVAFSADVTPV-----N 231
DB 223 ENATKLADSFVKQVLYKEKITGVEEAHN--KAQPNYSFVGYSDVITGTGOTSIGNWDY 281
QY 232 ARTVW--NGDEPSSRIANT-----NSITDVSNIWISLAGTNTKYOFSPNSYGPSTGY 282
DB 282 AORTITNGDEP--RSISNTPADGQTMVQPLSNVSIWISLAGTNGAKYTLTTEYIGSTGY 339
QY 283 LYEPYKLVKAAADANNVGLQYKLNNGVQVEATSTSA-----NNTTANPTPAVDEIKVA 337
DB 340 LYFPYKLVNTSDQMKLGLEYKLND-----ATEPSAITFGSEQTMNGKTPVNDINVA 391
QY 338 KIVLSGLRFGQNTIELSVPTGEGNMKNKVPAMIGNIYLSNNENADK1 384
DB 392 KVTLANLKFGSNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWKKI 433

RESULT 9

```
Q9ZHR9
AC Q9ZHR9; PRELIMINARY; PRT; 644 AA.
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE HYPOTHETICAL 69.9 KDA PROTEIN.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31, ATCC19610;
RC MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (M9) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmga
RT family."
RT Infect. Immun. 66:5570-5575(1998).
DR EMBL; AF053978; AAC69274.1; -.
KW Hypothetical protein.
SQ SEQUENCE 644 AA; 69866 MW; 8B9F352B13FBDE5C CRC64;

Query Match 37.2%; Score 745; DB 2; Length 644;
Best Local Similarity 42.3%; Pred. No. 7e-29;
Matches 172; Conservative 65; Mismatches 122; Indels 48; Gaps 11;

QY 9 DANPNQG-----TQLEAARMELTDLINAKAMTSLASLDQYAKIEASLSAYSAETVNN 61
DB 44 DTNPGDQGMNAAASQELAAARMGLTTVFDSKAKNLGLYVDYKKTQDTLTTRAYDAKTVL 103
QY 62 NNLNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
DB 104 DNSSSTTNLEAKRLETAIRTAATSKQTFDEQHAELVKVIEKLTLSNETATLAPYA 163
QY 122 STAYNQIRNVLVDLYNKASSLITKTLPLNGGTLSDSNEITTANKNINNTL--STINBQK 179
DB 164 AAQYAGIKMHLGSLYDAGKAITTKTLEPEVGDPLTASAVMANTKIVEAIKDEVLPKQ 222
QY 180 TNADALSNTFKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPV-----NKKY 231
DB 223 ENATKLADSFVKQVLVEKIKTGVVEAHNS-QPANYSFVGSVDITGTANGQTSIPNNY 281
QY 232 ARRTVW-NGDEPSSRLANT-----NSITDYSWIYSLAGTNTKYQFSNYPSTGY 282
DB 282 AORTIFTNGDEP--RSVSNTPVGDQTMQAPLSNVSWIYSLAGTAKYTFLEFYTGSTGY 339
QY 283 LFYPKLVKAADANNVGLQYKLNNGVQVFEATSTSA-----NNTTANPTPAVDEIKVA 337
DB 340 LFYPKLVNTSDQVKGLEYKLN-----ATKPSAITFGSEQTMNGKTPVNDINVA 391
QY 338 KIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNENNAADKI 384
DB 392 KVTLANFGSKIEFSVP-----VEKYSPIGNMNYLSSSPNNWNI 433

RESULT 10
Q49495
ID Q49495 PRELIMINARY; PRT; 650 AA.
AC Q49495; O08060;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE HYPOTHETICAL 69.9 KDA PROTEIN.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG31, ATCC19610;
RC MEDLINE=99003182; PubMed=9784576;
RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;
RT "A protein (M9) associated with monoclonal antibody-mediated
RT agglutination of Mycoplasma gallisepticum is a member of the pmga
RT family."
RT Infect. Immun. 66:5570-5575(1998).
DR EMBL; AF053978; AAC69274.1; -.
KW Hypothetical protein.
SQ SEQUENCE 644 AA; 69866 MW; 8B9F352B13FBDE5C CRC64;
```

```
RC STRAIN=S6;
RC MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum."
RN FEBS Lett. 352:347-352(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RC Markham P.F.;
RC Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U09714; AAB50152.1; -.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 650 HAEMAGGLUTININ
SQ SEQUENCE 650 AA; 70249 MW; 3ABACDB65940EBBB CRC64;

Query Match 37.1%; Score 744; DB 2; Length 650;
Best Local Similarity 42.9%; Pred. No. 7.9e-29;
Matches 172; Conservative 62; Mismatches 127; Indels 40; Gaps 11;

QY 11 NPNNGQ-----TQLEAARMELTDLINAKAMTSLASLDQYAKIEASLSAYSAETVNNN 63
DB 51 NPGDQGMNAAQELAAARMGLTTVFDSKAKNLGLYVDYKKTQNTLTTRAYDAKTVLDN 110
QY 64 LNATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLSST 123
DB 111 SSTTQNLNEAKRLETAIRTAATSKQTFDEQHAELVKVIEKLTLSNETATLAPYADA 170
QY 124 AYQIRNVLVDLYNKASSLITKTLPLNGGTLSDSNEITTANKNINNTL--STINBQKTN 181
DB 171 QYAGIKMHLGSLYDAGKAITTKTLEPEVGDPLTASAVMANTKIVEAIKDEVLPKKN 229
QY 182 ADALSNSFTKKVIQNNQSFVGTFTNANVQPSNYSFVAFSADVTPV-----NKKYAR 233
DB 230 ATKLADSFVKQVLVEKIKTGVVEAHN-KAPANYSFVGSVDITGTANGQTSIPNNYAQ 288
QY 234 RRTVW-NGDEPSSRLANT-----NSITDYSWIYSLAGTNTKYQFSNYPSTGYLY 284
DB 289 RRTIFTNGDEP--RSVSNTPVGDQTMQAPLSNVSWIYSLAGTAKYTFLEFYTGSTGYLY 346
QY 285 LFYPKLVKAADANNVGLQYKLNNGVQVFEATSTANNTTAN-PTPAVDEIKVAKTVLSG 343
DB 347 LFYPKLVNTSDQVKGLEYKLNDA----TSPSITFGSEQTMNGKTPVNDINVAKTVLAN 402
QY 344 LRFQNTIELSVPTGEGNNKVPAMIGNIYLSNENNAADKI 384
DB 403 LFYGSNKIEFSVPA-----EKVSPMIGNMNYLSSSPNNWNI 438

RESULT 11
Q49497
ID Q49497 PRELIMINARY; PRT; 649 AA.
AC Q49497;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE HYPOTHETICAL 69.9 KDA PROTEIN.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RC MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whithear K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, pmga, of Mycoplasma gallisepticum."
RN FEBS Lett. 352:347-352(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S6;
RC Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U09714; AAB50152.1; -.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 650 HAEMAGGLUTININ
SQ SEQUENCE 650 AA; 70249 MW; 3ABACDB65940EBBB CRC64;
```

```
RL FEBS Lett. 352:347-352(1994).
DR EMBL; L28424; AAA62416.1; -.
KW Signal.
FT SIGNAL
SQ SEQUENCE 649 AA; 70205 MW; 356554BD2C72C1F8 CRC64;

Query Match 37.1%; Score 743; DB 2; Length 649;
Best Local Similarity 42.3%; Pred. No. 8.8e-29;
Matches 172; Conservative 62; Mismatches 125; Indels 48; Gaps 11;

QY 9 DANPNNGQ-----TQLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEATVN 61
DB 48 DTNPGDGGGMMNAASQELAAARMGLTTTFDSKAKNLGLYDYKKTQNTLTTKAYDAKTVL 107
QY 62 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
DB 108 DNSSSTTONLNEAKTRLETAIRTAATSQTDEQHAELVKYKELKTTLSNETATLAPYA 167
QY 122 STAYNOIRNNLDVLYNKASSLITKTLPLNGTLLDSNEITTANKNINNTL--STINEOK 179
DB 168 DAQYAGIKHMLSGLYDAGKAITTKTLEPVEGDP-LTASAVMMANTKIWEAKDEVLNPK 226
QY 180 TNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPV-----NKY 231
DB 227 ENATKLADUSFVKQVLKKEITGVEEAHN-KAQPANISFVGISVDITGTTGTSIPNWDY 285
QY 232 ARRTYV-NGDEPSSRLANT-----NSTIDVSWIYSLAGTNTKYQFSFSGY 282
DB 286 AQTIFTSNDEP--RSISNTPADGOTMAQPLSNVSWIYSLAGTAKYLTFTFYGPSTGY 343
QY 283 LYFPYKLYKAADANNVGLYKLNNGVQOQEFATSTSA-----NNTTANPTPAVDEIKVA 337
DB 344 LYFPYKLVNTSDQVKGLEYKLD-----ATKPSAITFGSDOTMNGKTPVNDINVA 395
QY 338 KIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNENNADKI 384
DB 396 KYTLANLNGSNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 437

RESULT 12
ID Q49468 PRELIMINARY; PRT; 647 AA.
AC Q49468; Q53303;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HEMAGGLUTININ HOMOLOG PRECURSOR.
GN PMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molllicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93162830; PubMed-8432610;
RA Markham P.F., Giew M.D., Whithear K.G., Walker I.D.;
RT "Molecular cloning of a member of the gene family that encodes pmGA, a
hemagglutinin of Mycoplasma gallisepticum."
RL Infect. Immun. 61:903-909(1993).
DR EMBL; M83178; AAA02996.1; -.
DR EMBL; S55216; AAB25397.2; -.
KW Signal.
FT SIGNAL
FT CHAIN 1 25 POTENTIAL.
FT CHAIN 26 647 HEMAGGLUTININ HOMOLOG.
SQ SEQUENCE 647 AA; 70333 MW; 33916673BB9E28C4 CRC64;

Query Match 35.6%; Score 713; DB 2; Length 647;
Best Local Similarity 41.1%; Pred. No. 2.5e-27;
Matches 169; Conservative 60; Mismatches 124; Indels 59; Gaps 12;

QY 9 DANPNNGQ-----TQLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEATVN 61
DB 48 DTNPGDGGGMMNAASQELAAARMGLTTTFDSKAKNLGLYDYKKTQNTLTTKAYDAKTVL 107
QY 62 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
DB 108 DNSSSTTONLNEAKTRLETAIRTAATSQTDEQHAELVKYKELKTTLSNETATLAPYA 167
QY 122 STAYNOIRNNLDVLYNKASSLITKTLPLNGTLLDSNEITTANKNINNTL--STINEOK 179
DB 168 DAQYAGIKHMLSGLYDAGKAITTKTLEPVEGDP-LTASAVMMANTKIWEAKDEVLNPK 226
QY 180 TNADALSNSFIKKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPV-----NKY 231
DB 227 ENATKLADUSFVKQVLKKEITGVEEAHN-KAQPANISFVGISVDITGTTGTSIPNWDY 285
QY 232 ARRTYV-NGDEPSSRLANT-----NSTIDVSWIYSLAGTNTKYQFSFSGY 282
DB 286 AQTIFTSNDEP--RSISNTPADGOTMAQPLSNVSWIYSLAGTAKYLTFTFYGPSTGY 343
QY 283 LYFPYKLYKAADANNVGLYKLNNGVQOQEFATSTSA-----NNTTANPTPAVDEIKVA 337
DB 344 LYFPYKLVNTSDQVKGLEYKLD-----ATKPSAITFGSDOTMNGKTPVNDINVA 395
QY 338 KIVLSGLRFGQNTIELSVPTGEGNNKVPAMIGNIYLSNENNADKI 384
DB 396 KYTLANLNGSNKIEFSVPA-----EKVSPMIGNMYLSSSPNNWNI 437

RESULT 13
ID Q9KH15 PRELIMINARY; PRT; 656 AA.
AC Q9KH15;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ADHESIN PMGAL.2.
GN PMGAL.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molllicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS;
RA Shen Q.C., Bi D.R., Weng C.J.;
RT "Sequence analysis of the pmGA multigene family of Mycoplasma
gallisepticum strain HS."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275312; AAF91413.1; -.
SQ SEQUENCE 656 AA; 70875 MW; 58EA7E075FC617E1 CRC64;

Query Match 35.1%; Score 704; DB 2; Length 656;
Best Local Similarity 41.0%; Pred. No. 6.9e-27;
Matches 166; Conservative 67; Mismatches 124; Indels 48; Gaps 12;

QY 9 DANPNNGQ-----TQLEAARMELTDLINAKAMTLASLDQYAKIEASLSAYSEATVN 61
DB 63 DTNPGGQNMDSAAQELTAARTALTSLSKANVANVEMTSIAKTQNTLLIAYTTAEQTS 122
QY 62 NNLNATLEOLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 121
DB 123 QNSSATLEQVKNTSALQPAINTANSNKQKFDQDHSNLLMSYKNLMATLAKKETAVMTLK 182
QY 122 STAYNOIRNNLDVLYNKASSLITKTLPLNGTLLDSNEITTANKNINNTL--TINEOK 179
DB 183 DPKYSAILDQINGVSKGEELVQHTLDPVS-GIVPAANTITEITKIEVISEKTLQDOK 241
QY 180 TNADALSNSFI--KKVIONNEQSFVGTFTNANVOPSNYSFVAFSADVTPV----- 227
DB 242 NNAQDQFANYQSTFLDKTKLENVEDA-----KKMGOPANYSFVGISVDITGTSQETTIP 295
QY 228 NYKYARRTYWNGDEPSSRLANTNS-----ITDVSWIYSLAGTNTKYQFSFSGY 279
```

```
Db      296   NWNFAQRAITSGNQPTKATTTGGEDQGSTAKPLSDVSWIAGTGAKYTLLETFYGPS    355  
        |:|:|:: :|| :: ||:: | |:: ||||||| || |:||  
  
Qy     280   TGVLFPYPKLVAADANNYGLOVKLN-NGNVQQVEFATSTSANTNTANPTPAVDIKVAK    338  
        ||:||||||| |:|||||:: :| |:: | |:: |:  
  
Db     356   TGVLPFKLKVAND--DVGOLKLSNETLPIIFEGCT---TTNGPAAVTVENIVAK    409  
        ||:||||||| |:|||||:: :| |:: | |:: |:  
  
Qy     339   IVLSGLRFGONTIELSVPTGEGNNKWVPADMIGNILSSNENNADK    383  
        |:|:| ||||| || |:|||||:: |||||:: |:  
  
Db     410   VRUTGLAFGKNATIEFSVP-----MSKVAPMGNMYITSSDTETNK    449  
        |:|:| ||||| || |:|||||:: |||||:: |:  
  
RESULT 14  
O05122 PRELIMINARY; PRT; 703 AA.  
AC O05122;  
DC 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DI 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE HAEMAGGLUTININ.  
DN PMGAL_9.  
OS Mycoplasma gallisepticum.  
OC Bacterii; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI_TaxID=2096;  
ON [1]  
OR SEQUENCE FROM N.A.  
RP STRAIN=S6;  
RC MEDLINE=95010739; PubMed=7925999;  
RX Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,  
RA Browning G.F., Whithear K.G., Walker I.D.;  
RT "The organisation of the mycogene family which encodes the major cell  
RN surface protein, pmga, of Mycoplasma gallisepticum.";  
RL PBBS Lett. 352:347-352(1994).  
RS [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S6;  
RX Markham P.F.;  
RA Submitted (Feb-1997) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; U90714; AAB50154.1; -.  
DL InterPro: IPRO02819; HD.  
SQ SEQUENCE 703 AA; 75742 MW; 310B6BE9F73CBC5 CRC64;
```

Query Match 34.5%; Score 690.5; DB 2; Length 703;
Best Local Similarity 35.0%; Pred. No. 3.4e-26;
Matches 167; Conservative 75; Mismatches 120; Indels 115; Gaps 12;

Qy	2	CMSITKDKANP-----NNGQT-----QLEARMELTDLINAK	33
Db	26	CTSATIPNPTNPBPKPDPMPNPBPGMGNGNTNPGMDTAAQELASAKAALTTLTNRE	85
Qy	34	AMTILASLDQYAKIEASLSAYSEAEVTNNLNATLEQLKMAKTNIESAINQANTDKTTFD	93
Db	86	SEKVGLYVDYAKIKADLTSAYTVAKTSDSSTSLVQVKTATSKLQTAIDKAASDKQFE	145
Qy	94	NEHPNLVEAYKALKTTLEOR-ATNLEGLSSTAYNOIRNNLVLDLYNKASSLITKTLPLNG	152
Db	146	QDHKDLMPYSEUKTTLQKRNATVL--LQQPKYSALINKINSIYAQGEVWIRTLDPVVS-	202
Qy	153	GTLLDSNEITTAANKNNITLS--TINEQRTNADALSNGFKKVIQNNQSFVGTETNANV	210
Db	203	GAIPTAASITKVDEINKAISENLQPKKDNADAFANYQFFKL---DKTKIMGNSTNMK	259
Qy	211	QPSNYSEVAFPSADVTV-----NYKARTVTVMNGDEBPSSRIILANTNS-----	252
Db	260	QPONYSFVGYSVGVTGMSQCGTTIPNWNFAQIRVSSGAPRAPLASQETPOAETPPMSA	319
Qy	253	-----IDTVSWIYSIA	263
Db	320	POGVEPAQQGDSSPKOASETQEVSTPAAEVQAQADTEQPATSGGTPLTIDTVSWISLS	379
Qy	264	GTNTKIQFSFSPNYGPGSTGYLYFPYKLVKAADANNVGLQYKLANGNVQOVEATISANNT	323

```

Db      380 GTDVKYFTTFNYFGPSMAYLYFPYKLVKSD--SVGLQYKLNNNPVALNFGSETNAN-- 435
        || : || | : ||| ||||| : | : ||||| : | : | : |||
Qy      324 TANPTPAVDKEIKVAKIVSLGRLTGONTIELSVPTGEGNMKNKVAPMIGNIYLSNENN 380
        || || |||| : | : ||||| ||| ||||| : ||| : | : |||
Db      436 -GPAASVDNIINAKVLNLANFGEINTIEFSVP-----MNKVAPMIGNMWITSDVAN 485
        || || |||| : | : ||||| ||| ||||| : ||| : | : |||

RESULT 15
Q9L8D5 PRELIMINARY; PRT; 419 AA.
ID Q9L8D5
AC AC
DT DT Q9L8D5;
DT 01-OCT-2000 (TrEMBLrel_15, Created)
DT 01-OCT-2000 (TrEMBLrel_15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel_15, Last annotation update)
DE DE PGA-LIKE PROTEIN 9.3 (FRAGMENT).
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_Taxid=2096;
RN [.]
RP SEQUENCE FROM N.A.
RC STRAIN=F;
RA Pharr G.T., Branton S.L., Hanson L.A., Minion F.C., Lott B.D.,
RA May J.D., Hughlett M.B.;
RT "A novel pmGA-like gene from the F-strain (vaccine strain) of
RT Mycoplasma gallisepticum."
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF210770; AAF29525.1; -.
SQ NON_TER 419 419
SQ SEQUENCE 419 AA; 45072 MW; FE5BE37FDDB3B0C CRC64;
```

Query Match 30.9%; Score 619.5; DB 2; Length 419;
Best Local Similarity 40.1%; Pred. No. 5.4e-23;
Matches 149; Conservative 58; Mismatches 122; Indels 43; Gaps 11;

[illegible]

Search completed: June 12, 2002, 10:50:42
Job time: 205 sec

